THE UNIVERSITY *of* EDINBURGH

Simple Models for Resolving Environments in Disordered Alloys by X-ray Photoelectron Spectroscopy

Thomas Underwood

Doctor of Philosophy
The University of Edinburgh
2012

Abstract

In disordered alloys, atoms belonging to the same chemical element will exhibit different environments. This leads to variations in the atoms' local electronic structures, which in turn leads to variations in the binding energies of their core levels. These binding energies can be measured experimentally using core level X-ray photoelectron spectroscopy (XPS). Therefore, in theory at least, core level XPS can be used to resolve different environments in alloys. However, to make this a reality one must understand how an atom's local electronic structure, and hence the binding energies of its core levels, are affected by local environment. In this thesis, two simple phenomenological models are explored which purport to correctly describe the local electronic structure of disordered alloys. The first model which we consider has its roots in chemical intuition; specifically, the notion that pairs of unlike atoms, i.e. atoms belonging to different chemical elements, transfer a certain quantity of charge, while like atoms do not. Using this model - known as the optimised linear charge model (OLCM) - the relationship between an atom's local electronic structure, core level binding energies, and its environment is explored in detail, both in the bulk of disordered alloys and near their surfaces. As well as 'homogeneous' disordered alloys, in which the concentrations of the alloy's constituent elements are the same throughout the entire alloy, various 'inhomogeneous' disordered alloy systems are considered. These include alloys exhibiting surface segregation - in which the concentrations at the surface differ from those in the bulk - as well as interfaces between two metals with various levels of intermixing. The results of our investigation of bulk inhomogeneous alloys are compared to analogous *ab initio* results, which confirms the model's viability as a tool for rationalising the relationship between local electronic structure, core level binding energies, and environment. More generally, our results also reveal a number of interesting new phenomena. Firstly, the widths of spectra in inhomogeneous disordered alloys are significantly larger in some cases than is possible in any analogous homogeneous disordered alloy. Secondly, differences between the concentrations of each element at the surface and deep within the bulk cause a shift in the work function of the alloy under consideration. The latter results in qualitatively different trends than one

would expect if this phenomenon was ignored, and prompts an alternative interpretation of the results of a recent experimental study. The second model which we consider is a particular case of the charge-excess functional model, in which the realised charges on all atoms are those which minimise a particular expression for the total energy of the system, and whose accuracy has been well established. The underlying assumptions and properties of this model are explored in detail, adding insight into the nature of the screening and inter-atomic interactions in disordered alloys. The model is shown to be equivalent to the OLCM for the case of binary alloys, and can therefore be considered to be the generalisation of the OLCM for alloys containing more than two chemical elements. The model is also used to derive analytical expressions for various physical quantities for any alloy, including the width of core level XPS spectra and the Madelung energy. These expressions are then used to investigate how the physical quantities to which they pertain vary with the concentrations of each element in a homogeneous disordered alloy consisting of three elements. Among other things, it was observed that the width of the core level XPS spectra is maximised when the concentrations of the two elements in the alloy with the largest electronegativity difference have equal concentrations, while the remaining element has a vanishing concentration.

Declaration

Except where otherwise stated, the research undertaken in this thesis was the unaided work of the author. Where the work was done in collaboration with others, a significant contribution was made by the author.

T. Underwood
July 2012

Acknowledgements

I would like to thank all of those who have helped me in the course of this project. In particular, I would like to thank my supervisor Jamie Cole, not only for giving me the opportunity to carry out this project, but also for his excellent supervision throughout it and ongoing support as it draws to a close. His thorough proof reading of this thesis and subsequent comments are particularly appreciated. I also gratefully acknowledge EPSRC, who funded this project.

Contents

List of Figures

List of Tables

List of Abbreviations

CEFM Charge-excess functional model

CLS Core level shift

DFT Density functional theory

ETF Embedded thin film

FWHM Full width at half maximum

GLCM Generalised linear charge model

HAU Hartree atomic units

LCD Localised charge distribution

LCM Linear charge model

LDA Local density approximation

LIZ Local interaction zone

LSGF Locally self-consistent Green's function

LSMS Locally self-consistent multiple scattering

MLCM Multi-shell linear charge model

OLCM Optimised linear charge model

RS Ruban and Skriver

SSCPA Single site coherent potential approximation

SSLSGF Single site locally self-consistent Green's function

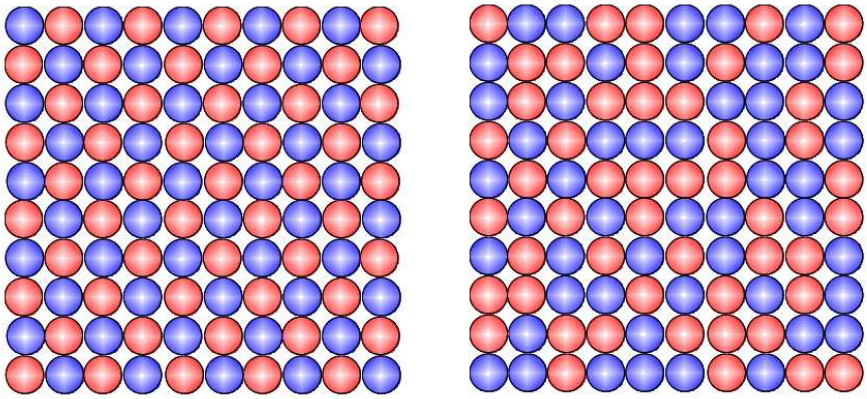
XPS X-ray photoelectron spectroscopy

Chapter 1

Introduction

Alloys are metallic mixtures of two or more elements, at least one of which is a metal. They have been one of the most important materials since ancient times, and are ubiquitous in the modern world. There is a continual demand for new alloys with desired physical properties (e.g. high strength, radiation resistance), and the design and development of such alloys drives forward technological advances. However, success in this area is inextricably linked to how well we understand their underlying physics.

Substitutional alloys are a large class of alloys which include bronze (a mixture of copper and tin) and brass (a mixture of copper and zinc). Below their melting point, substitutional alloys consist of one or more distinct substitutional solid solution phases, in which the positions of the atoms are idealised as forming a periodic crystal lattice, though the pattern formed by considering the *species* of the different atoms is not necessarily periodic - where by the 'species' of an atom we mean the element to which it belongs. If the pattern *is* periodic, then the phase is referred to simply as an *ordered alloy*; otherwise, it is referred to as a *disordered alloy*. Schematic illustrations of both are given in Fig. 1.1. In this thesis, we examine disordered alloys. These present a challenge to describe theoretically on account of their lack of periodicity. However, it is of fundamental importance to understand their properties: the phase diagrams of most substitutional alloys are dominated by disordered alloy phases. In a disordered alloy, the atoms belonging to the same species exhibit a wide range of different environments. This can be seen in Fig. 1.1(b): there are 'red' atoms with 4, 3, 2, 1 or 0 'unlike' (i.e. 'blue') nearest neighbour atoms, and similarly for the blue atoms. Since the electronic structure 'at' a particular atom depends in part on the atom's environment, we expect that the electronic structure at different atoms belonging to the same species within a disordered alloy will vary widely. Unfortunately, the exact nature of the relationship between an atom's environment and its electronic structure in such alloys is not currently well-understood.



(a) Ordered alloy (b) Disordered alloy

Figure 1.1: A schematic illustration of an ordered alloy and a disordered alloy. Here, each circle represents an atom, and the colour of the circle reflects the species of the atom.

The aim of this thesis is to make progress in understanding this relationship. We will focus particularly on the relationship between an atom's environment and the binding energies of its *core levels*, which are those electronic energy levels which are so tightly bound to the atom's nucleus that they do not take part in chemical bonding. A quantitative understanding of this would enable core level X-ray photoelectron spectroscopy (XPS), which can measure these binding energies experimentally, to be used to determine the proportions of different environments present in a given alloy. Such 'environment-resolved spectroscopy' would be a useful technique for the atomic-scale characterisation of alloys - an area which is becoming increasingly important with the advent of nanotechnology. Specifically, the above aim will be accomplished through the use of 'simple' phenomenological models; as opposed to through the use of *ab initio* methods, which determine the electronic structure using only the principles of quantum mechanics, and have no 'free parameters' whose values must be determined by other means, e.g. experiment. Such models are useful for a variety of reasons. If they are accurate, then they can provide an elegant description of a system, as well as being used as a faster alternative to computationally expensive *ab initio* calculations, or as a simple framework for analysing experimental results. Even qualitatively accurate models can provide useful information. For example, they can provide insight into how a number of factors which cannot be easily 'separated out' from *ab initio* calculations influence the system under consideration. They can also be used to predict qualitative trends.

1.1 Thesis Layout

The outline of the rest of this thesis is as follows:

- Chapter 2 contains the relevant background theory to this thesis. We begin this chapter with a general discussion of electronic structure. We then review the key findings of *ab initio* calculations with regards to the atomic-scale electronic structure of disordered alloys, and describe various simple models which have been proposed to explain these results. Finally, we discuss core level XPS, and review the relevant studies pertaining to disordered alloys.

- Chapter 3 considers the *optimised linear charge model*, which is a simple phenomenological model for determining atomic-scale electronic structure. The model is used to examine the atomic-scale electronic structure and the distribution of core level binding energies in a large variety of alloy systems. These include disordered alloys in which the concentration of each species is *inhomogeneous* throughout the alloy, as well as the surface regions of such alloys.

- Chapter 4 considers another simple phenomenological model - the *generalised linear charge model*. Firstly, the properties of this model are explored in detail. The model is then used to explore how various physical quantities in *ternary* disordered alloys, which consist of three species, depend on the concentrations of these species.

- Chapter 5 contains a summary of the key findings of this thesis and suggests topics for future study.

Chapter 2

Background Theory

In this chapter we present the background theory relevant to this thesis.

2.1 Electronic structure

To begin, we will consider the behaviour of a general non-relativistic quantum system of N_{elec} interacting electrons in the presence of an 'external' potential

$$V_{\text{ext}}(\mathbf{r}) = -\sum_i \frac{1}{4\pi\epsilon_0} \frac{e^2 z_i}{|\mathbf{r} - \mathbf{R}_i|} \tag{2.1}$$

due to the N_{nuc} atomic nuclei, where the index i labels the atomic nuclei, z_i is the atomic number of nucleus i, $\mathbf{R}_i$ is the position of nucleus i, e is the fundamental charge, and ϵ_0 is the vacuum permittivity.[1] The state of the system is determined by the many-electron wavefunction $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_{\text{elec}}})$, where $\mathbf{x}_m = (\mathbf{r}_m, s_m)$ is the combined position and spin coordinates pertaining to electron m, and s_m can be one of two values: $\uparrow$ or $\downarrow$. Since we are considering a system of indistinguishable fermions, Ψ must be antisymmetric under exchange of any pair of electron labels, i.e.

$$\Psi(\ldots, \mathbf{x}_m, \ldots, \mathbf{x}_n, \ldots) = -\Psi(\ldots, \mathbf{x}_n, \ldots, \mathbf{x}_m, \ldots). \tag{2.2}$$

Furthermore, it must be correctly normalised:

$$\langle \Psi | \Psi \rangle = 1, \tag{2.3}$$

[1] Note that $V_{\text{ext}}(\mathbf{r})$ is negative because the electronic energy is lowered in the presence of atomic nuclei.

where

$$\langle \Psi | \Phi \rangle = \int d\mathbf{x}_1 \dots \int d\mathbf{x}_{N_{\text{elec}}} \Psi^*(\mathbf{x}_1, \dots, \mathbf{x}_{N_{\text{elec}}}) \Phi(\mathbf{x}_1, \dots, \mathbf{x}_{N_{\text{elec}}}), \tag{2.4}$$

and the notation $\int d\mathbf{x}$ denotes an integral over all of position-spin space, i.e.

$$\int d\mathbf{x} = \sum_{s=\uparrow,\downarrow} \int d\mathbf{r}. \tag{2.5}$$

The (expected) value O of any observable for a state Ψ can be found by using the equation

$$O = \langle \Psi | \hat{O} \Psi \rangle, \tag{2.6}$$

where $\hat{O}$ is the Hermitian operator corresponding to the observable. The most important observable for our purposes is the electronic energy E. This is the sum of the observables E_{kin}, E_{ext} and E_{int}, i.e.

$$E = E_{\text{kin}} + E_{\text{ext}} + E_{\text{int}}, \tag{2.7}$$

where E_{kin} is the kinetic energy of the electrons, E_{ext} is the *external energy*, which is the energy stored in the electron-nucleus interactions, and E_{int} is the *internal energy*, which is the energy stored in the electron-electron interactions. Respectively, the operators associated with E, E_{kin}, E_{ext} and E_{int} are

$$\hat{\mathcal{H}} = \hat{\mathcal{T}} + \hat{\mathcal{V}}_{\text{ext}} + \hat{\mathcal{V}}_{\text{int}}, \tag{2.8}$$

$$\hat{\mathcal{T}} = -\frac{\hbar^2}{2m_e} \sum_m \nabla_m^2, \tag{2.9}$$

$$\hat{\mathcal{V}}_{\text{ext}} = \sum_m V_{\text{ext}}(\mathbf{r}_m) \tag{2.10}$$

and

$$\hat{\mathcal{V}}_{\text{int}} = \frac{1}{2} \sum_m \sum_{n \neq m} \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_m - \mathbf{r}_n|}, \tag{2.11}$$

where the indices m and n label the electrons and take the possible values $1, 2, \dots, N_{\text{elec}}$, ∇_m^2 is the Laplace operator with respect to the coordinates $\mathbf{r}_m$, $\hbar$ is the reduced Planck constant, and m_e is the electron mass. For the sake of brevity we will henceforth adopt Hartree atomic units (HAU), in which $\hbar = m_e = e = 4\pi\epsilon_0 = 1$, unless otherwise stated. The units of the relevant physical quantities to this thesis in HAU are given in Table 2.1.

Determining the *energy eigenstates* is the primary aim of electronic structure theory. These are the wavefunctions which, in addition to satisfying said constraints, correspond

Quantity	Expression	Value
Angular momentum	$\hbar$	4.13567×10^{-15} eV s
Mass	m_e	9.10938×10^{-31} kg
Charge	e	1.60218×10^{-19} C
Length	$a_0 = 4\pi\epsilon_0\hbar^2/(m_e e^2)$	0.529177 Å
Energy	Ha $= e^2/(4\pi\epsilon_0 a_0)$	27.2114 eV
Electric potential	Ha$/e$	27.2114 V

Table 2.1: The units of measurement of various physical quantities in HAU. $\hbar$ is the reduced Planck's constant, m_e is the electron mass, e is the elementary charge, ϵ_0 is the permittivity of free space, a_0 is the Bohr radius, and Ha is the Hartree energy. Values are only quoted to 6 significant figures.

to stationary points in the energy E. Of particular interest is the *ground state* Ψ_0, which is the energy eigenstate with the lowest energy. Unfortunately, due to the complexity of Eqn. (2.8), it is impossible to determine the energy eigenstates directly. We must simplify the problem of determining the energy eigenstates somehow.

2.1.1 The Hartree-Fock Approximation

It is instructive to consider the *Hartree-Fock approximation*. In this, Ψ is assumed to have the form

$$\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_{N_\text{elec}}) = \frac{1}{\sqrt{N_\text{elec}!}}\begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_1(\mathbf{x}_2) & \ldots & \psi_1(\mathbf{x}_{N_\text{elec}}) \\ \psi_2(\mathbf{x}_1) & \psi_2(\mathbf{x}_2) & \ldots & \psi_2(\mathbf{x}_{N_\text{elec}}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{N_\text{elec}}(\mathbf{x}_1) & \psi_{N_\text{elec}}(\mathbf{x}_2) & \ldots & \psi_{N_\text{elec}}(\mathbf{x}_{N_\text{elec}}) \end{vmatrix}, \tag{2.12}$$

where the set of functions $\psi_1(\mathbf{x}), \psi_2(\mathbf{x}), \ldots, \psi_{N_\text{elec}}(\mathbf{x})$ form an orthonormal set, i.e.

$$(\psi_p|\psi_q) = \delta_{pq}, \tag{2.13}$$

where[2]

$$(\psi|\phi) = \int d\mathbf{x}\; \psi^*(\mathbf{x})\phi(\mathbf{x}), \tag{2.14}$$

and the indices p and q label these functions and have the possible values $1, 2, \ldots, N_\text{elec}$. The functions $\psi_1(\mathbf{x}), \psi_2(\mathbf{x}), \ldots, \psi_{N_\text{elec}}(\mathbf{x})$ describe the behaviour of a single electron, and are known as *spin orbitals*. The form of Ψ in Eqn. (2.12) is the simplest which

[2]Note that $(\psi|\phi)$ denotes the inner product between two *one-electron* wavefunctions $\psi(\mathbf{x})$ and $\phi(\mathbf{x})$, and contains a single integral over all of position-spin space, while $\langle\Psi|\Phi\rangle$ denotes the inner product between two *many-electron* wavefunctions $\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_{N_\text{elec}})$ and $\Phi(\mathbf{x}_1,\ldots,\mathbf{x}_{N_\text{elec}})$, and contains multiple integrals over all of position-spin space (see Eqn. (2.4)).

satisfies Eqns. (2.3) and (2.2). Because of this, the Hartree-Fock approximation is useful for elucidating all of the qualitative features present in 'real' electronic structures.

The physical significance of assuming that Ψ has the form given in Eqn. (2.12) is best illustrated by expanding Eqn. (2.12) for small values of N_{elec}. For $N_{\text{elec}} = 2$, Eqn. (2.12) becomes

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2!}}\Big[\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2) - \psi_1(\mathbf{x}_2)\psi_2(\mathbf{x}_1)\Big]. \tag{2.15}$$

In the above, the first term corresponds to a state in which electron 1 occupies spin orbital 1, and electron 2 occupies spin orbital 2; and the second term corresponds to a state in which electron 2 occupies spin orbital 1, and electron 1 occupies spin orbital 2. Thus the above Ψ is the superposition of all possible states of the system in which spin orbitals 1 and 2 are occupied by the 2 electrons. Note that the second term is negative, which ensures that Eqn. (2.2) is obeyed, and that the factor of $1/\sqrt{2!}$, in conjunction with Eqn. (2.13), ensures that Eqn. (2.3) is obeyed. For $N_{\text{elec}} = 3$, Eqn. (2.12) becomes

$$\begin{aligned}\Psi(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \frac{1}{\sqrt{3!}}\Big[&\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)\psi_3(\mathbf{x}_3) - \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_3)\psi_3(\mathbf{x}_2) \\ &- \psi_1(\mathbf{x}_2)\psi_2(\mathbf{x}_1)\psi_3(\mathbf{x}_3) + \psi_1(\mathbf{x}_2)\psi_2(\mathbf{x}_3)\psi_3(\mathbf{x}_1) \\ &+ \psi_1(\mathbf{x}_3)\psi_2(\mathbf{x}_1)\psi_3(\mathbf{x}_2) - \psi_1(\mathbf{x}_3)\psi_3(\mathbf{x}_2)\psi_3(\mathbf{x}_1)\Big].\end{aligned} \tag{2.16}$$

Here, the first term in the above equation corresponds to a state in which electron 1 occupies spin orbital 1, electron 2 occupies spin orbital 2 and electron 3 occupies spin orbital 3; the second term corresponds to a state in which electron 1 occupies spin orbital 1, electron 3 occupies spin orbital 2 and electron 2 occupies spin orbital 3; etc. In this case, Ψ is the linear combination of all possible states of the system in which the spin-orbitals 1, 2 and 3 are occupied by the 3 electrons. Again, the signs of the terms ensure that Eqn. (2.2) is obeyed, and the factor of $1/\sqrt{3!}$ ensures that Eqn. (2.3) is obeyed. The above examples illustrate that, in the Hartree-Fock approximation, Ψ is a superposition of all possible states in which the spin orbitals $1, 2, \ldots, N_{\text{elec}}$ are occupied by the N_{elec} electrons.

Before determining the energy eigenstates, it is important to know how electrons are *correlated* in the Hartree-Fock approximation, i.e. how the probability of finding an electron at a particular position $\mathbf{r}$ depends on the particular positions of all other electrons. This can be determined by calculating the quantities $n(\mathbf{r}, s)$ and $n(\mathbf{r}, s, \mathbf{r}', s')$, where $n(\mathbf{r}, s)$ is the number of electrons with spin s at position $\mathbf{r}$, and $n(\mathbf{r}, s, \mathbf{r}', s')$ is the number of pairs of electrons with spins s and s' located at positions $\mathbf{r}$ and $\mathbf{r}'$ respectively.

The operators corresponding to these quantities are

$$\hat{n}(\mathbf{r}, s) = \sum_m \delta(\mathbf{r} - \mathbf{r}_m)\delta_{ss_m} \tag{2.17}$$

and

$$\hat{n}(\mathbf{r}, s, \mathbf{r}', s') = \sum_m \sum_{n \neq m} \delta(\mathbf{r} - \mathbf{r}_m)\delta_{ss_m}\delta(\mathbf{r}' - \mathbf{r}_n)\delta_{s's_n}, \tag{2.18}$$

where $\delta(\mathbf{r})$ is the Dirac delta function and $\delta_{ss'}$ is the Kronecker delta for a pair of spins s and s'. Applying Eqn. (2.6), and assuming that Eqn. (2.2) holds, $n(\mathbf{r}, s)$ and $n(\mathbf{r}, s, \mathbf{r}', s')$ are given by

$$n(\mathbf{r}, s) = N_{\text{elec}} \int d\mathbf{x}_2 \int d\mathbf{x}_3 \ldots \int d\mathbf{x}_{N_{\text{elec}}} |\Psi\big((\mathbf{r}, s), \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_{N_{\text{elec}}}\big)|^2 \tag{2.19}$$

and [1]

$$n(\mathbf{r}, s, \mathbf{r}', s') = N_{\text{elec}}(N_{\text{elec}} - 1) \int d\mathbf{x}_3 \ldots \int d\mathbf{x}_{N_{\text{elec}}} |\Psi\big((\mathbf{r}, s), (\mathbf{r}', s'), \mathbf{x}_3, \ldots, \mathbf{x}_{N_{\text{elec}}}\big)|^2 \tag{2.20}$$

in general. It can be shown that, in the Hartree-Fock approximation, these quantities become [1]

$$n(\mathbf{r}, s) = \sum_p |\psi_p(\mathbf{r}, s)|^2 \tag{2.21}$$

and

$$n(\mathbf{r}, s, \mathbf{r}', s') = n(\mathbf{r}, s)n(\mathbf{r}', s') - \delta_{ss'} \left| \sum_p \psi_p^*(\mathbf{r}, s)\psi_p(\mathbf{r}', s) \right|^2. \tag{2.22}$$

If the electrons were uncorrelated, we would expect that

$$n(\mathbf{r}, s, \mathbf{r}', s') = n(\mathbf{r}, s)n(\mathbf{r}', s'). \tag{2.23}$$

As can be seen from Eqn. (2.22), this is the case if $s' \neq s$, i.e. pairs of electrons with unlike spins are uncorrelated; however, it is not the case for pairs of electrons with the same spin. For electrons with the same spin, $n(\mathbf{r}, s, \mathbf{r}', s')$ is lower than would be expected if the electrons were uncorrelated. In fact, the second term in Eqn. (2.22) becomes equal to the first if $\mathbf{r}' = \mathbf{r}$ and $s' = s$, and hence $n(\mathbf{r}, s, \mathbf{r}, s) = 0$. In other words, it is impossible for two electrons with the same spin to be found at the same position. This phenomenon is known as the *exchange hole*: electrons with the same spin are correlated such that they 'repel' each other. It should be emphasised that the aforementioned correlation between electrons is not due to a force in the conventional sense: it follows solely from the fact that Ψ has the form given in Eqn.

(2.12). Specifically, it follows from the requirement that Ψ satisfies Eqn. (2.2).

Let us now attempt to determine the energy eigenstates within the Hartree-Fock approximation. First we must obtain expressions for $E_{\rm kin}$, $E_{\rm ext}$ and $E_{\rm int}$. Let us begin with the kinetic energy $E_{\rm kin}$. Noting that $E_{\rm kin} = \langle\Psi|\hat{\mathcal{T}}\Psi\rangle$ (see Eqn. (2.6)), we find that $E_{\rm kin}$ is simply the sum of the kinetic energies of each occupied spin orbital, i.e.

$$E_{\rm kin} = \sum_p (\psi_p|\Big[-\frac{1}{2}\nabla^2\Big]\psi_p). \tag{2.24}$$

Similarly, for the external energy we find that

$$E_{\rm ext} = \int d\mathbf{r}\; V_{\rm ext}(\mathbf{r}) n(\mathbf{r}), \tag{2.25}$$

where the *electron density* $n(\mathbf{r})$ is given by

$$n(\mathbf{r}) = \sum_{s=\uparrow,\downarrow} n(\mathbf{r}, s) = \sum_p \sum_{s=\uparrow,\downarrow} |\psi_p(\mathbf{r}, s)|^2. \tag{2.26}$$

The internal energy is more interesting. For this, we find that

$$E_{\rm int} = E_{\rm H} + E_{\rm x}, \tag{2.27}$$

where

$$E_{\rm H} = \frac{1}{2}\int d\mathbf{r} \int d\mathbf{r}'\; \frac{1}{|\mathbf{r}-\mathbf{r}'|} n(\mathbf{r}) n(\mathbf{r}') \tag{2.28}$$

is the classical electrostatic interaction energy between the electrons - known as the *Hartree energy*; and

$$E_{\rm x} = \frac{1}{2}\sum_p (\psi_p|\hat{V}_{\rm x}\psi_p) \tag{2.29}$$

is known as the *exchange energy*, where the the operator $\hat{V}_{\rm x}$ is defined by the equation

$$\hat{V}_{\rm x} f(\mathbf{x}) = -\sum_q \psi_q(\mathbf{x}) \int d\mathbf{x}' \frac{1}{|\mathbf{r}-\mathbf{r}'|} \psi_q^*(\mathbf{x}') f(\mathbf{x}') \tag{2.30}$$

with $f(\mathbf{x})$ an arbitrary function. The exchange energy has no classical analogue. It arises from the phenomenon of the exchange hole described above, and is negative, i.e. the exchange hole phenomenon acts to lower the internal energy of the electrons. This makes sense: because of the exchange hole, on average there are fewer electrons in the vicinity of any particular electron than if there were no exchange hole, and hence the (positive) Coulomb repulsive energy between the electrons is lowered.

Substituting the above equations into Eqn. (2.7) gives an expression for E in the

Hartree-Fock approximation. This is [2]

$$\begin{aligned} E = &\sum_p \int d\mathbf{x}\ \psi_p^*(\mathbf{x})\big(-\frac{1}{2}\nabla^2\big)\psi_p(\mathbf{x}) + \sum_p \int d\mathbf{x}\ V_{\text{ext}}(\mathbf{r})|\psi_p(\mathbf{x})|^2 \\ &+ \frac{1}{2}\sum_p\sum_q \int d\mathbf{x} \int d\mathbf{x}'\ \frac{1}{|\mathbf{r}-\mathbf{r}'|}|\psi_p(\mathbf{x})|^2|\psi_q(\mathbf{x}')|^2 \\ &- \frac{1}{2}\sum_p\sum_q \int d\mathbf{x} \int d\mathbf{x}'\ \psi_p^*(\mathbf{x})\psi_q(\mathbf{x})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_p(\mathbf{x}')\psi_q^*(\mathbf{x}'). \end{aligned} \tag{2.31}$$

The spin orbitals which make this stationary give the Ψ (via Eqn. (2.12)) which correspond to energy eigenstates. Using the method of Lagrange multipliers to ensure that the constraints described by Eqn. (2.13) are adhered to, one finds that the spin orbitals corresponding to stationary points in the energy E obey the following set of N_{elec} equations [2]:

$$\hat{F}\psi_p = \sum_q \varepsilon_{pq}\psi_q, \tag{2.32}$$

where

$$\hat{F} = -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + \hat{V}_{\text{x}} \tag{2.33}$$

is known as the the *Fock operator*,

$$V_{\text{H}}(\mathbf{r}) = \int d\mathbf{r}'\ \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.34}$$

is the electrostatic potential at $\mathbf{r}$ due to all the electrons in the system - known as the *Hartree potential*, and the values of the Lagrange multipliers ε_{pq} are as yet unknown, but obey the relation $\varepsilon_{pq} = \varepsilon_{qp}^*$. The set of equations described by Eqn. (2.32) are known as the *Hartree-Fock equations*. These can be simplified by appealing to the fact that there is some flexibility in the choice of the spin orbitals: there are many possible sets of functions $\psi_1, \psi_2, \ldots, \psi_{N_{\text{elec}}}$ which give the same Ψ in Eqn. (2.12). It can be shown (see, e.g. Refs. [2,3] for more details) that one can choose the spin orbitals such that ε_{pq} is non-zero only if $q = p$, in which case the Hartree-Fock equations simplify to the *canonical Hartree-Fock equations*:

$$\hat{F}\psi_p = \varepsilon_p\psi_p, \tag{2.35}$$

where $\varepsilon_p = \varepsilon_{pp}$. A set of spin orbitals $\psi_1, \psi_2, \ldots, \psi_{N_{\text{elec}}}$ and their corresponding values $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{N_{\text{elec}}}$ which are a solution to these equations, when substituted into Eqn.

(2.12), give the wavefunction for an energy eigenstate of the system whose energy is [2]

$$\begin{aligned} E =& \sum_p \varepsilon_p - \frac{1}{2}\sum_p \sum_q \int d\mathbf{x} \int d\mathbf{x}' \; \frac{1}{|\mathbf{r}-\mathbf{r}'|}|\psi_p(\mathbf{x})|^2|\psi_q(\mathbf{x}')|^2 \\ &+ \frac{1}{2}\sum_p \sum_q \int d\mathbf{x} \int d\mathbf{x}' \; \psi_p^*(\mathbf{x})\psi_q(\mathbf{x})\frac{1}{|\mathbf{r}-\mathbf{r}'|}\psi_p(\mathbf{x}')\psi_q^*(\mathbf{x}') \\ =& \sum_p \varepsilon_p - E_{\mathrm{H}} - E_{\mathrm{x}}. \end{aligned} \tag{2.36}$$

The above equation follows from combining Eqns. (2.35) and (2.31).

As can be seen from Eqn. (2.33), the Fock operator can be considered to be a Hamiltonian describing a one-electron system in which the potential operator is $\hat{V} = V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + \hat{V}_{\mathrm{x}}$. Eqn. (2.35) therefore reveals that, in an energy eigenstate of the entire system, spin orbital p is an energy eigenstate of the aforementioned one-electron system with energy eigenvalue ε_p. In other words, an electron occupying any particular spin orbital 'experiences' the nuclei through the potential $V_{\mathrm{ext}}(\mathbf{r})$, and the other electrons in the system through the Hartree potential $V_{\mathrm{H}}(\mathbf{r})$ and the (rather enigmatic) operator $\hat{V}_{\mathrm{x}}$. Note that, as can be seen from Eqn. (2.36), the spin orbital eigenvalues do not sum to the total energy E. ε_p cannot therefore be interpreted as the energy of an electron in spin orbital p. Its actual physical significance is as follows: ε_p is the negative of the energy required to remove an electron from spin orbital p whilst leaving the electrons in the other spin orbitals unaffected. This is *Koopmans' theorem*, which will be referred to later.

Solving the canonical Hartree-Fock equations is not straightforward. The difficulty originates from the fact that the Fock operator depends on all of the occupied spin orbitals,[3] as well as the fact that the Fock operator contains both differential and integral operators. An analytical solution can be obtained only in the special case of the *homogeneous electron gas*, in which $V_{\mathrm{ext}}(\mathbf{r})$, and hence $n(\mathbf{r})$, is constant everywhere. This is derived in many texts (e.g. Refs. [1, 4]). Here we only quote the result for the exchange energy per unit volume, which we will use later. This is

$$\epsilon_{\mathrm{x,HEG}}(n) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} n^{4/3}, \tag{2.37}$$

where n is the density of the homogeneous electron gas. Numerical methods can be used to solve the canonical Hartree-Fock equations for free atoms, though these methods cannot be applied to molecules and solids [3]. For molecules and solids, one must resort

[3] This can be seen from Eqns. (2.33), (2.34) and (2.30): the Fock operator $\hat{F}$ depends on $V_{\mathrm{H}}(\mathbf{r})$ and $\hat{V}_{\mathrm{x}}$, which themselves depend on all of the occupied spin orbitals.

to further approximations, such as the Roothan and Slater approximations (for details, see, for instance, Ref. [5]).

As was shown earlier, electrons with opposite spins are uncorrelated in the Hartree-Fock approximation. This means that $n(\mathbf{r}, s, \mathbf{r}, s') = n(\mathbf{r}, s)n(\mathbf{r}, s') \neq 0$, i.e. it is possible to find two electrons with opposite spins at the same position $\mathbf{r}$. This is manifestly not the case in reality: the Coulomb repulsion between electrons would prevent this from happening. The preceding discussion highlights the shortcoming of the Hartree-Fock approximation: it does not account for the correlation between the electrons due to Coulomb repulsion; though it still captures the most important features of electronic structure. Nowadays, a plethora of different methods exist in order to calculate the energy eigenstates which somehow incorporate the effects of 'Coulomb correlation' between electrons. Such methods are known collectively as *post-Hartree-Fock methods* (see Ref. [6]). Although they provide greater accuracy than methods based upon the Hartree-Fock approximation, post-Hartree-Fock methods are extremely computationally demanding, and their use is limited to small systems, i.e. they cannot be used for solids. The only methods which are capable of incorporating the effects of Coulomb correlation for solids are those based upon *density functional theory*, which we will now discuss.

2.1.2 Density Functional Theory

Density functional theory (DFT) is based upon the two Hohenberg-Kohn theorems [7]. Proofs of these theorems, which we are about to state, along with details of other subtleties regarding the theorems can be found in Refs. [1, 2]. The first Hohenberg-Kohn theorem is:

Theorem 2.1. *The external potential $V_{ext}(\mathbf{r})$ is determined uniquely by the ground state electron density $n_0(\mathbf{r})$.*[4]

In other words, there is a map from a particular ground state density $n_0(\mathbf{r})$ to a single form of external potential $V_{\text{ext}}(\mathbf{r})$. Consider now Eqns. (2.8), (2.9), (2.10) and (2.11). Note that N_{elec} is determined by $n_0(\mathbf{r})$:

$$N_{\text{elec}} = \int d\mathbf{r}\; n_0(\mathbf{r}). \tag{2.38}$$

With this and the above theorem in mind, and noting that the summations over m

[4]Strictly, $V_{\text{ext}}(\mathbf{r})$ is only determined uniquely, *except for an additive constant*, by the ground state electron density $n_0(\mathbf{r})$. However, we will deliberately ignore this detail in subsequent discussions. This is justified because, as is well known, the addition of an arbitrary constant to the potential of a system does not alter its underlying physics. The effect is to alter the zero of the absolute energy scale, which is unimportant since only *differences* in energy, and not absolute energies, have a physical significance.

and n run from 1 to N_{elec}, it follows that $\hat{\mathcal{H}}$ must be uniquely determined by $n_0(\mathbf{r})$. Furthermore, since all physical properties of the system depend on $\hat{\mathcal{H}}$, it follows that *all physical properties of the system are uniquely determined by* $n_0(\mathbf{r})$. For each physical property P therefore there exists a universal functional[5] $P[n(\mathbf{r})]$ which gives the value of P for a system in which $n(\mathbf{r})$ is the ground state density, where by the term 'universal' in this context we mean 'the same for all systems'. Of particular importance are the ground state kinetic energy E_{kin0}, the ground state Hartree energy E_{H0}, and the ground state *exchange-correlation energy* E_{xc0}. The exchange-correlation energy E_{xc} is defined by the equation

$$E_{\text{xc}} = E_{\text{int}} - E_{\text{H}}, \tag{2.39}$$

i.e. it is the contribution to the internal energy other than the Hartree energy. This brings us to the second Hohenberg-Kohn theorem:

Theorem 2.2. *For a system whose external potential is* $V_{ext}(\mathbf{r})$*, the functional* $E_{HK}[n(\mathbf{r})]$*, which is defined as*

$$E_{HK}[n(\mathbf{r})] = E_{kin0}[n(\mathbf{r})] + E_{H0}[n(\mathbf{r})] + E_{xc0}[n(\mathbf{r})] + \int d\mathbf{r}\, V_{ext}(\mathbf{r})n(\mathbf{r}), \tag{2.40}$$

is at a global minimum at the system's ground state density $n_0(\mathbf{r})$ *given the constraints on* $n(\mathbf{r})$ *that* $n(\mathbf{r}) \geq 0$ *and* $\int d\mathbf{r}\, n(\mathbf{r}) = N_{elec}$*. Furthermore, the value of this functional at* $n_0(\mathbf{r})$ *is the system's ground state energy* E_0*, i.e.* $E_0 = E_{HK}[n_0(\mathbf{r})]$*.*

This implies that E_0 and $n_0(\mathbf{r})$ for a given system can be determined by minimising the functional $E_{\text{HK}}[n(\mathbf{r})]$ with respect to $n(\mathbf{r})$. To put this into practice, however, we must know the functionals $E_{\text{kin0}}[n(\mathbf{r})]$, $E_{\text{H0}}[n(\mathbf{r})]$ and $E_{\text{xc0}}[n(\mathbf{r})]$. The second of these is already known: by definition, $E_{\text{H0}}[n(\mathbf{r})]$ is equal to the right-hand side of Eqn. (2.28). Unfortunately, the functionals $E_{\text{kin0}}[n(\mathbf{r})]$ and $E_{\text{xc0}}[n(\mathbf{r})]$ are *not* known.

The fact that we do not know $E_{\text{kin0}}[n(\mathbf{r})]$ can be sidestepped by assuming that that the many-electron wavefunction Ψ pertaining to the electron density $n(\mathbf{r})$ has the same form as used in the Hartree-Fock approximation (Eqn. (2.12)), except that the set of coordinates $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_{\text{elec}}}$ describe a set of 'electron-like particles' instead of actual electrons. These particles are known as *quasielectrons*, and the aforementioned assumption is known as the *Kohn-Sham ansatz* [8]. In this case, Eqn. (2.26) holds. Furthermore, $E_{\text{kin0}}[n(\mathbf{r})]$ is given by the right-hand side of Eqn. (2.24). We therefore have an expression for $E_{\text{kin0}}[n(\mathbf{r})]$, not in terms of $n(\mathbf{r})$, but in terms of the quasielectron spin orbitals $\psi_1, \psi_2, \ldots, \psi_{N_{\text{elec}}}$. With this in mind, and using the aforementioned form

[5] A functional is a map from a function to a scalar value. If a quantity F is a functional of the function $f(x)$, then this is denoted $F[f(x)]$.

for $E_{\mathrm{H0}}[n(\mathbf{r})]$, Eqn. (2.40) becomes

$$E_{\mathrm{HK}}[n(\mathbf{r})] = \sum_p (\psi_p | \Big[-\frac{1}{2}\nabla^2 \Big] \psi_p) + \frac{1}{2} \int d\mathbf{r} \int d\mathbf{r}' \, \frac{1}{|\mathbf{r} - \mathbf{r}'|} n(\mathbf{r}) n(\mathbf{r}') + E_{\mathrm{xc0}}[n(\mathbf{r})] + \int d\mathbf{r} \, V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}), \tag{2.41}$$

where $n(\mathbf{r})$ is related to the spin orbitals by Eqn. (2.26). Recall that we wish to minimise $E_{\mathrm{HK}}[n(\mathbf{r})]$ with respect to $n(\mathbf{r})$. Given that $n(\mathbf{r})$ depends entirely on the quasielectron spin orbitals, this is equivalent to minimising $E_{\mathrm{HK}}[n(\mathbf{r})]$ with respect to all of the spin orbitals. Using a similar procedure to that used to obtain the canonical Hartree-Fock equations in Section 2.1.1, it can be shown that the spin orbitals which correspond to stationary points in $E_{\mathrm{HK}}[n(\mathbf{r})]$ obey a set of N_{elec} equations - known as the *Kohn-Sham equations*. These are

$$\hat{H} \psi_p = \varepsilon_p \psi_p, \tag{2.42}$$

where

$$\hat{H} = -\frac{1}{2}\nabla^2 + U(\mathbf{r}), \tag{2.43}$$

$$U(\mathbf{r}) = V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}), \tag{2.44}$$

$$V_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc0}}}{\delta n(\mathbf{r})} \tag{2.45}$$

is the *exchange-correlation potential*, and $\delta E_{\mathrm{xc0}} / \delta n(\mathbf{r})$ denotes the functional derivative of $E_{\mathrm{xc0}}[n(\mathbf{r})]$ with respect to $n(\mathbf{r})$.[6] Similarly to the Hartree potential $V_{\mathrm{H}}(\mathbf{r})$ (see Eqn. (2.28)), the exchange-correlation potential is a functional of $n(\mathbf{r})$, though we do not know its form since we do not know the form of $E_{\mathrm{xc0}}[n(\mathbf{r})]$. We will ignore this fact for a moment, and proceed assuming that the form of $E_{\mathrm{xc0}}[n(\mathbf{r})]$ *is* known.

Note that, as can be seen from Eqn. (2.42), the quasielectron spin orbitals corresponding to stationary points in $E_{\mathrm{HK}}[n(\mathbf{r})]$ are energy eigenstates of a one-electron system whose Hamiltonian operator is $\hat{H}$. $\hat{H}$ therefore plays the same role for the quasielectron spin orbitals as the Fock operator does for the electron spin orbitals in

[6] The functional derivative $\delta F / \delta f(x)$ of the functional $F[f(x)]$ is defined by the equation

$$F[f(x) + \delta f(x)] - F[f(x)] = \int dx \, \frac{\delta F}{\delta f(x)} \delta f(x),$$

where $\delta f(x)$ is an infinitesimally small function. E.g. if $F[f(x)] = \int dx \, f(x)^2$, then

$$F[f(x) + \delta f(x)] - F[f(x)] = \int dx \, 2 f(x) \delta f(x),$$

where we have ignored terms $O(\delta f(x)^2)$ since $\delta f(x)$ is small. Comparing this to the definition of $\delta F / \delta f(x)$ above, we find that $\delta F / \delta f(x) = 2 f(x)$ in this case.

the Hartree-Fock approximation discussed in Section 2.1.1. However, in this case there is no complicated exchange operator $\hat{V}_{\mathrm{x}}$ in the Hamiltonian, only the density-dependent potential $U(\mathbf{r})$. This makes the Kohn-Sham equations far easier to solve than the canonical Hartree-Fock equations. Furthermore, the effects of Coulomb correlation, which are ignored in the Hartree-Fock approximation, are included in the Kohn-Sham approach.

A set of quasielectron spin orbitals which solve the Kohn-Sham equations correspond to a stationary point in the functional $E_{\mathrm{HK}}[n(\mathbf{r})]$ with energy [1]

$$E_{\mathrm{HK}}[n(\mathbf{r})] = \sum_p \varepsilon_p - E_{\mathrm{H0}}[n(\mathbf{r})] + E_{\mathrm{xc0}}[n(\mathbf{r})] - \int d\mathbf{r}\; V_{\mathrm{xc}}(\mathbf{r}) n(\mathbf{r}), \tag{2.46}$$

which follows from combining Eqn. (2.42) (in conjunction with Eqns. (2.43), (2.44) and (2.45)) and Eqn. (2.41). Such a stationary point, which we hope is the global minimum in $E_{\mathrm{HK}}[n(\mathbf{r})]$ [9], can be found by using the following iterative procedure:

1. Make a guess $n^{\mathrm{in}}(\mathbf{r})$ for the ground state electron density.

2. Use $n^{\mathrm{in}}(\mathbf{r})$ to calculate $U(\mathbf{r})$. The relevant equations are Eqns. (2.44), (2.34) and (2.45). This $U(\mathbf{r})$ gives a guess for $\hat{H}$ via Eqn. (2.43).

3. Calculate the eigenstates of $\hat{H}$ and their corresponding eigenvalues. In other words, solve the Kohn-Sham equations (Eqn. (2.42)) pertaining to $\hat{H}$.

4. The aforementioned eigenstates are all possibilities for the N_{elec} quasielectron spin orbitals $\psi_1, \psi_2, \ldots, \psi_{N_{\mathrm{elec}}}$. Choose the spin orbitals such that they are all *different* eigenstates. This is necessary to avoid a trivial solution: from Eqn. (2.12), it can be seen that two identical spin orbitals results in two identical rows of the determinant, which in turn results in the determinant, and hence Ψ, becoming 0.

5. With these spin orbitals, calculate the associated electron density $n^{\mathrm{out}}(\mathbf{r})$ by using Eqn. (2.26).

6. Check whether $n^{\mathrm{out}}(\mathbf{r}) = n^{\mathrm{in}}(\mathbf{r})$. If so, then we have a self-consistent solution to the Kohn-Sham equations, i.e. we have found a stationary point in $E_{\mathrm{HK}}[n(\mathbf{r})]$. If not, use $n^{\mathrm{in}}(\mathbf{r})$ and $n^{\mathrm{out}}(\mathbf{r})$ to formulate a better guess for the ground state electron density than we had originally, and repeat the whole procedure until self-consistency is reached.

In step 3, the eigenstates and eigenvalues of $\hat{H}$ can be determined using a plethora of different methods and approximations, an overview of which is given in Ref. [1].

Of particular relevance to this thesis is the *KKR-Green's function method*, which is particularly suited to disordered alloys. In this method, the system is first split into regions associated with each nucleus. Then one determines how each region individually scatters an incoming electronic wave into an outgoing wave. (All electron waves are time-independent here). This requires knowledge of the one-electron potential within each region, which in the above procedure is provided by the $U(\mathbf{r})$ calculated in step 2. The electronic stationary states for the whole system can then be determined from knowledge of the individual scattering properties of each site: they occur when, for all regions, the incoming electronic wave for that region is equal to the sum of outgoing electronic waves from all other regions. More information on the KKR-Green's function method can be found in Refs. [10–12].

As mentioned earlier, we do not know the form of $E_{\mathrm{xc0}}[n(\mathbf{r})]$, which is necessary in order to determine the form of $V_{\mathrm{xc}}(\mathbf{r})$, and ultimately enable the above procedure to determine the ground state of the system to be used in practice. This is where we must resort to approximations. Various approximations for $E_{\mathrm{xc0}}[n(\mathbf{r})]$ exist. In the *local density approximation* (LDA) $E_{\mathrm{xc0}}[n(\mathbf{r})]$ is assumed to have the form

$$E_{\mathrm{xc0}}[n(\mathbf{r})] = \int d\mathbf{r}\ \epsilon_{\mathrm{x,HEG}}(n(\mathbf{r})) + \int d\mathbf{r}\ \epsilon_{\mathrm{c,HEG}}(n(\mathbf{r})), \tag{2.47}$$

where $\epsilon_{\mathrm{x,HEG}}(n)$ and $\epsilon_{\mathrm{c,HEG}}(n)$ are the exchange energy and the *correlation energy* per unit volume for a homogeneous electron gas of density n respectively. Recall that the exchange energy is defined as the non-classical contribution to the total energy within the Hartree-Fock approximation. The correlation energy E_{c} is defined as the difference between the total non-relativistic energy of the system under consideration and that predicted within the Hartree-Fock approximation. To clarify, the correlation energy is the *correction* to the Hartree-Fock energy; and its name reflects the fact that a correction is required because Coulomb correlation between electrons is ignored in the Hartree-Fock approximation. Note that we have an analytical expression for $\epsilon_{\mathrm{x,HEG}}(n)$ (Eqn. (2.37)). Unfortunately, no such expression exists for $\epsilon_{\mathrm{c,HEG}}(n)$, except for the limits $n \to \infty$ and $n \to 0$. However, various accurate analytical expressions for $\epsilon_{\mathrm{c,HEG}}(n)$ have been derived from fitting curves to results from numerical calculations of the correlation energy for the homogeneous electron gas at various densities, and these have found widespread use. (See Ref. [1] for more details on parameterisations of $\epsilon_{\mathrm{c,HEG}}(n)$). The LDA is obviously exact for the homogeneous electron gas, and is expected to be valid for systems in which the electron density varies slowly with position. However, we have no *a priori* reason to suppose that it will be useful in systems where the electron density varies rapidly with position, such as atoms and molecules. Despite this, the LDA performs remarkably well when applied to such

systems: ionisation energies of atoms, dissociation energies of molecules, and cohesive energies calculated using the LDA have an accuracy of typically 10-20%; and bond lengths have an accuracy of $\sim$ 1% [13]. In cases where the LDA does not meet the required level of accuracy, one can use more sophisticated approximations for $E_{\mathrm{xc0}}[n(\mathbf{r})]$. These include the *local spin density approximation*, which is similar to the LDA except that one separately keeps track of the electron density associated with $\uparrow$ and $\downarrow$ electrons; and the class of *generalised gradient approximations*, in which the gradient of the electron density at each position is taken into account when calculating $E_{\mathrm{xc0}}[n(\mathbf{r})]$. Of course, with the gain in accuracy associated with more sophisticated approximations comes a greater computational cost. Further discussion on the topic of various approximations for $E_{\mathrm{xc0}}[n(\mathbf{r})]$ can be found in Ref. [1].

Having outlined how to calculate the electronic ground state, we will now conclude this section by discussing the physical significance of the quasielectron energy eigenvalues. Recall that in the Hartree-Fock approximation the significance of the eigenvalues is given by Koopmans' theorem: assuming that spin orbitals are unaffected by the removal of any one electron from the system, the values of $-\varepsilon_p$ are the energies required to remove individual electrons from the system, i.e. they are the possible ionisation energies of the system. This result does *not* apply to the Kohn-Sham eigenvalues, since they do not pertain to actual electrons, but to quasielectrons. The only exception to this is the highest energy Kohn-Sham eigenvalue in a finite system, which is the negative of the ionisation energy of the system's most weakly bound electron [1]. In spite of the above, one often regards the Kohn-Sham eigenvalues as describing actual electrons, in which case Koopmans' theorem will apply. This should be done with caution; the validity of such an approach hinges on whether the quasielectron spin orbitals are a good representation of the actual electron spin orbitals (as would be obtained from a perfect post-Hartree-Fock calculation). A mathematical significance to the Kohn-Sham eigenvalues is provided by *Janak's theorem* [14]:

$$\varepsilon_p = \frac{\partial E_0}{\partial \eta_p}, \tag{2.48}$$

where η_p, which takes a value between 0 and 1, is the number of electrons occupying spin orbital p; and E_0 is the ground state total energy of the system.[7] We will use this theorem later.

[7]For the sake of brevity, we have been sloppy in our statement of Janak's theorem. Strictly, E_o in Eqn. (2.48) should be replaced with an energy which is rigorously defined for all possible sets of values of η_p, which is not the case for E_0. See Ref. [14] for details.

2.2 Local Electronic Structure of Disordered Alloys

In the previous section we described the features of electronic structure and outlined how DFT can be used to determine the electronic ground state of a nonspecific system. In this section, we consider the electronic structure of disordered alloys. Specifically, we consider how it varies locally, where by 'locally' in this context we mean on the scale of the distance between adjacent nuclei. Given the complexity of disordered alloys, it is necessary to make some simplifying assumptions in order to convert the problem of determining their electronic structure into a tractable one. We will now describe some such assumptions which will be used throughout the rest of this thesis. As alluded to in Chapter 1, we will assume that the positions of the nuclei in the alloy lie upon an undistorted crystal lattice, whose lattice type and lattice parameter are known. (The geometry of the crystal lattice can be determined for a particular alloy, for example, by X-ray crystallography). It should be emphasised that in a 'real' disordered alloy this would not be the case: the differences in the atomic sizes of the constituent species in the alloy would cause the nuclear positions to form a 'distorted' crystal lattice. We hope that the distortions are small enough to be of no consequence. With this first assumption, we can specify an alloy by giving the underlying crystal lattice (i.e. the lattice type and lattice parameter) and the species of each nucleus, where by the *species* of a nucleus we are referring to its atomic number.[8] This assumption also allows us to partition space into identical Wigner-Seitz cells centred on each nucleus.[9] We will refer to the Wigner-Seitz cell centred on nucleus i and all of its contents collectively as *site* i. Denoting the volume of an individual site as V_{WS}, we will further assume that each site is spherical with radius R_{WS}, where R_{WS} - known as the *Wigner-Seitz radius* - is defined by the equation

$$V_{\mathrm{WS}} = \frac{4}{3}\pi R_{\mathrm{WS}}^3. \tag{2.49}$$

This assumption is reasonable if the lattice type for the system under consideration has a high packing fraction - which is almost always the case in metallic systems. Our final assumption regards the form of the one-electron potential $U(\mathbf{r})$. We will assume that $U(\mathbf{r})$ within site i is spherically symmetric about $\mathbf{R}_i$. Thus we can express $U(\mathbf{r})$ as

$$U(\mathbf{r}) = \sum_i u_i(|\mathbf{r} - \mathbf{R}_i|), \tag{2.50}$$

[8] Note that N_{elec} is specified if the species of all nuclei are known, since (assuming the system is charge neutral) $N_{\mathrm{elec}} = \sum_i z_i$.

[9] The Wigner-Seitz cell centred on a particular nucleus is defined as the set of points which are closer to that particular nucleus than any other.

where $u_i(r)$ describes the one-electron potential associated with site i, which is 0 if $r > R_{\mathrm{WS}}$. We will refer to this as the *spherical approximation*. A consequence of the spherical approximation is that the charge distribution within each site is spherically symmetric about the site's nucleus. We will exploit this fact later.

In addition to the above assumptions, it is common to idealise the disordered alloy as a *random alloy*. Let c_X denote the fraction of nuclei in the alloy which belong to species X. We will refer to c_X as the *concentration* of species X. Obviously, the concentrations of all N_{spec} species present in the alloy must obey

$$\sum_X c_X = 1, \tag{2.51}$$

where the summation over X is over the N_{spec} species. In a random alloy, the species of each nucleus is determined completely at random, and does not depend on the location of the nucleus itself, or on the species of any other nucleus. It should be emphasised that not all disordered alloys are random alloys. It is possible, for instance, to assign the species of nuclei in an alloy such that the X nuclei have on average more Y nuclei as nearest neighbours than would be case if the species of all nuclei were assigned completely at random, while still retaining the lack of periodicity characteristic of a disordered alloy. An alloy in which this is the case is said to exhibit a degree of *short range order*, which, by definition, is absent in random alloys.

2.2.1 The Single Site Approximation

Even with the above simplification to the one-electron potential $U(\mathbf{r})$, the problem of determining a given disordered alloy's electronic structure directly was until relatively recently regarded as intractable. An elegant way around this is provided by the *single site approximation*. This can be used to determine the average properties of the whole alloy, as well as the average properties of all sites belonging to a particular species. It does this through the construction of an *effective species* whose properties somehow reflect those of the alloy as a whole. The average properties of the whole alloy are then given by the properties of an *effective medium* which consists entirely of the effective species on the same underlying crystal lattice as the real alloy; and the average properties over all X sites are given by the properties of an X impurity embedded in the effective medium. This is illustrated in Fig. 2.1. The benefit of this approach is that the effective medium, and single impurities in the effective medium, are systems of sufficient simplicity to enable their electronic structures to be determined. The effective medium can be constructed in a number of ways. The *single site coherent potential approximation* (SSCPA) [15–17] is the most widely utilised method on account of its

accuracy and efficiency. In this, the effective medium is defined by the condition

$$\hat{G}_{\text{eff}}(\varepsilon) = \sum_X c_X \hat{G}_X(\varepsilon), \tag{2.52}$$

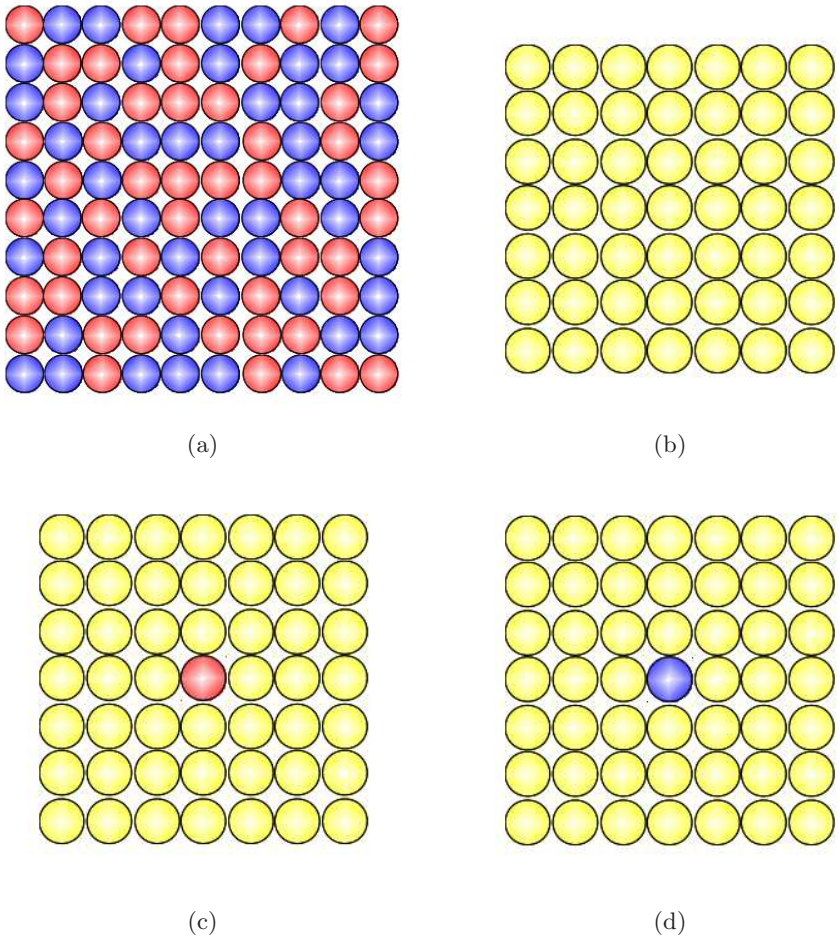
where

$$\hat{G}(\varepsilon) = (\varepsilon - \hat{H})^{-1} \tag{2.53}$$

is the *Green's function* for a system with one-electron Hamiltonian $\hat{H}$, $\hat{G}_{\text{eff}}(\varepsilon)$ is the Green's function for the pure effective medium, and $\hat{G}_X(\varepsilon)$ is the Green's function for the effective medium with a single X impurity. Note that the single site approximation does not provide information regarding the *fluctuations* in the properties of X sites which arise due to differences in their environments; it only provides the average properties of these sites. The single site approximation is therefore unsuitable for studying local electronic structure in a detailed manner. More information on single site approximations and the SSCPA is given in Refs. [10, 18].

2.2.2 Order-N Calculations

An approximation which *can* be used to investigate the details of local electronic structure in disordered alloys is to treat the alloy as a periodic system, where the repeating unit is a *supercell* containing many hundreds of sites whose species are assigned randomly - in a manner which reflects the nature of the substitutional disorder present in the disordered alloy. The larger the size of the supercell; the better this approximation becomes. However, such an approach could only be put into practice for the first time in the 1990s, with the development of so-called *order-N methods* in conjunction with improvements in computer hardware. The name of order-N methods reflects the fact that the computational effort required to determine the electronic structure of a system of N atoms scales linearly with N. By contrast, conventional DFT-based methods scale with N^3. Two order-N methods have been used to determine the electronic structure of random alloys. These are the locally self-consistent multiple scattering (LSMS) method [19], and the locally self-consistent Green's function (LSGF) method [20, 21]. Both are based upon the KKR-Green's function method described in Section 2.1.2. In the former, during step 3 in the DFT procedure for finding the electronic ground state described in Section 2.1.2, the properties of each site i are obtained by only considering electron wave scattering from the one-electron potentials on sites which are within a finite distance of site i. The region enclosed by this distance is known as a *local interaction zone* (LIZ). The properties obtained in this manner include the electron density within site i, which, along with the density from all other sites, is fed into step 5 in the procedure. Convergence of the electronic structure



(a) (b)

(c) (d)

Figure 2.1: An illustration of the single site approximation applied to a random alloy. The alloy under consideration is shown in (a), and consists of two species A and B which are represented by red and blue circles respectively. Yellow circles represent the effective species. The average properties of the whole alloy are determined from the effective medium, which is shown in (b); the average properties of A sites are determined from an A impurity embedded in the effective medium, which is shown in (c); and the average properties of B sites are determined from a B impurity embedded in the effective medium, which is shown in (d).

with the chosen size of the LIZs can be checked to ensure correct results. The LSGF method is similar to the LSMS method, except that, when determining the properties of site i, the region beyond the LIZ is an effective medium - instead of a vacuum as is the case in the LSMS. The benefit of this approach is that it allows the LIZ to be smaller, which reduces computational effort. In practice, the effective medium used in LSGF calculations is constructed using the same condition as that of the SSCPA, but if each site in the supercell were considered to belong to a different species. Specifically, consider the N 'impurity systems' formed by embedding each of the N sites in the supercell within the LSGF effective medium: the LSGF effective medium is defined such that the average Green's function of all of these impurity systems is equal to the Green's function of the pure effective medium.

2.2.2.1 The *Q-V* Relations

Order-N methods applied to disordered alloys have revealed the existence of the so-called *Q-V relations*, which is the observation that, for all sites belonging to the same species in a particular alloy, the *Madelung potential* - which we will define in a moment - of each site is a linear function of its total charge [22–24]. This holds over a wide range of disordered alloys, with various combinations of constituent species, concentrations of each species, and various degrees of substitutional disorder. (It should be pointed out that the spherical approximation was used in all the order-N calculations which we refer to). The Madelung potential is the electrostatic potential felt at site i's position due the charge distribution in all *other* sites. Given the spherical approximation, the Madelung potential for site i, which we will denote as V_i, is given by

$$V_i = \sum_j M_{ij} Q_j, \tag{2.54}$$

where M is the *Madelung matrix*, whose elements are defined as

$$M_{ij} = \begin{cases} 0 & \text{if } j = i, \\ 1/|\mathbf{R}_i - \mathbf{R}_j| & \text{otherwise.} \end{cases} \tag{2.55}$$

Eqn. (2.54) states that V_i is equivalent to the electrostatic potential at $\mathbf{R}_i$ if the total charge on all other sites was situated at their corresponding nuclear positions. This follows from the fact that the potential outside any spherically symmetric charge distribution is unchanged if all of the charge in the distribution was brought to its centre - a result which itself can be derived from Gauss' law. The *Q-V* relations can

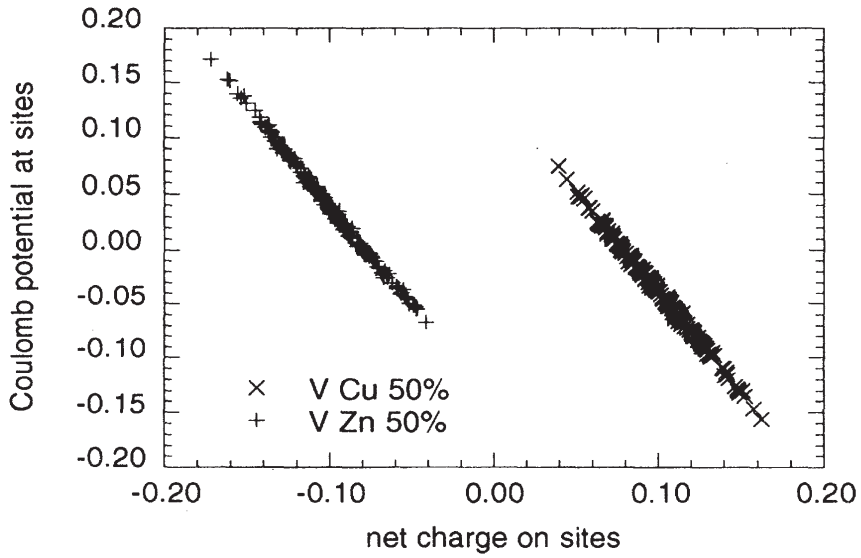


Figure 2.2: V_i (referred to as the Coulomb potential in the figure) vs. Q_i for sites in a $Cu_{0.5}Zn_{0.5}$ random alloy, as calculated using the LSMS method for a supercell containing 432 sites. Plus signs are data points for Zn sites, and crosses are for Cu sites. The quantities shown in the figure are in Rydberg atomic units, in which energies are measured in Rydbergs (2Ry=1Ha), distances in Bohr radii, and charges in units of e. (Taken from Ref. [22].)

therefore be expressed as

$$V_i = -a_i Q_i + k_i, \tag{2.56}$$

where a_i and k_i take the values a_X and k_X respectively if site i belongs to species X. Remarkably, the values of a_X and k_X for each species are ostensibly the same for all alloys, even ordered, which have the same species concentrations and underlying crystal lattice [22]. An example of the Q-V relations is shown in Fig. 2.2. This figure illustrates that the Q-V relations are not an exact result, though they do hold to a high degree of accuracy.

2.2.2.2 Universal Screening

More recently, Ruban and Skriver (RS) [24] discovered that there is a universality to the Q-V relations: for X sites in any alloy, the gradient a_X[10] and intercept k_X of their

[10]Throughout this thesis we will refer to the values of a_X as the 'gradients' of the Q-V relations; though strictly, as can be seen from Eqn. (2.56), they are the negative of the gradients of the Q-V relations.

corresponding Q-V relations obey

$$a_X R_{\mathrm{WS}} \approx 1.6 \tag{2.57}$$

and

$$\frac{k_X}{Q_X^{\mathrm{SSCPA}}} R_{\mathrm{WS}} \approx 1.6 \tag{2.58}$$

respectively, where Q_X^{SSCPA} is the charge of an X site obtained from a conventional SSCPA calculation for the given alloy. Strictly this universality is not exact, but holds to a good approximation, which is reflected in our use of the symbol '$\approx$' instead of '$=$' in the above two equations. Furthermore, RS utilised the *single site LSGF* (SSLSGF) method, in which the LIZ consists of only one site, and hence each site 'sees' the LSGF effective medium as its surroundings. Since a LIZ consisting of one site will not be 'converged' in the sense described earlier, results determined using the SSLSGF method are therefore not 'exact', though they are still at least semiquantitatively accurate. RS attributed the above result to a universal mechanism of screening, which we will now describe. Consider what happens if we perturb Q_i by an amount δQ_0, and allow the charge on all other sites to relax as to minimise the total energy of the system. Firstly, RS discovered that the induced change in charge for all sites in the βth nearest neighbour shell of i[11] are the same, which we will denote as δQ_β. Let the induced change in charge for a site in the βth nearest neighbour shell of i be denoted as δQ_β, and let δQ_β normalised to δQ_0 be denoted ϕ_β, i.e.

$$\phi_\beta = \frac{\delta Q_\beta}{\delta Q_0}. \tag{2.59}$$

RS also discovered that, regardless of the choice of site and the particular alloy,

$$\phi_\beta = u(R_\beta / R_{\mathrm{WS}}), \tag{2.60}$$

where R_β is the distance from site i to its βth nearest neighbour shell, and u is a universal function which is the same for all alloys. In other words, the values of ϕ_β obtained from considering all sites in all alloys, when plotted against $R_\beta / R_{\mathrm{WS}}$, lie upon a single curve. This curve is shown in Fig. 2.3. It is not at all obvious that this screening mechanism should lead to a universality in the Q-V relations. This is something which we will investigate in Chapter 4.

[11] The nearest neighbour shell of site i is the set of sites, not including site i itself, which are closest to i. The second nearest neighbour shell is the set of sites, not including site i itself or those sites in the nearest neighbour shell, which are closest to i. The βth nearest neighbour shell is the set of sites, not including site i or the $\beta, \beta-1, \ldots, 2, 1$th nearest neighbour shells, which are closest to i.

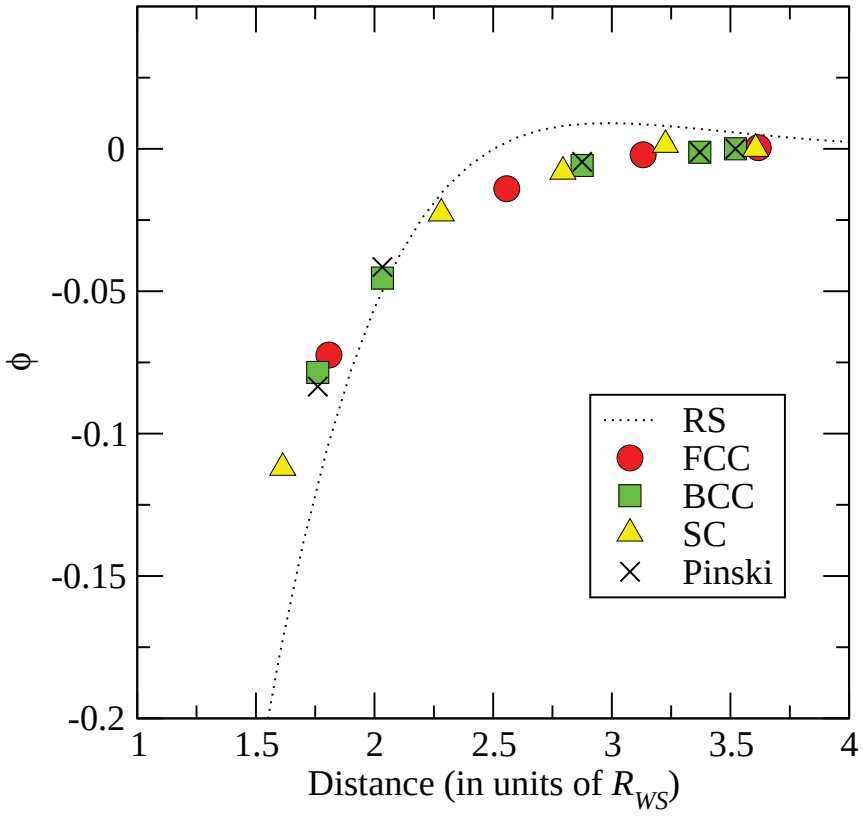


Figure 2.3: The distribution of screening charge for various models. The dotted curve traces the universal curve discovered by RS (Ref. [24]) utilising the SSLSGF method; and the crosses represent the curve obtained by Pinski's model (Ref. [25]). All other symbols represent the results of the OLCM (Ref. [26]). Circles, squares and triangles represent the OLCM results for fcc, bcc and sc lattices respectively. (Taken from Ref. [27].)

2.2.3 Simple Models

The existence of the Q-V relations and universal screening implies that a 'simple' model could possibly be created to describe the local electronic structure in disordered alloys. Such models are very much worth pursuing since they can act as a faster alternative to the computationally expensive order-N methods, aid in the interpretation of experimental results, and illuminate the underlying physics of complicated phenomena. We will now discuss several simple models which have been used in the past.

2.2.3.1 Pinski's Model

We will begin with a model used by Pinski in Ref. [25] in order to explore the nature of the Coulomb interactions in random alloys. This model is DFT-based, but with two major simplifications. Firstly, space is divided into cubes of dimension $\approx 1/4$ of the distance between adjacent nuclei, and within each of these cubes the electronic density constrained to be constant. Secondly, the *Thomas-Fermi approximation* is used in calculating the electronic kinetic energy. In the Thomas-Fermi approximation the kinetic energy functional $E_{\mathrm{kin0}}[n(\mathbf{r})]$ is assumed to be

$$E_{\mathrm{kin0}}[n(\mathbf{r})] = \int d\mathbf{r}\ \epsilon_{\mathrm{kin,HEG}}(n(\mathbf{r})), \tag{2.61}$$

where

$$\epsilon_{\mathrm{kin,HEG}}(n) = \frac{3}{10}(3\pi^2)^{2/3} n^{5/3} \tag{2.62}$$

is the kinetic energy per unit volume for a homogeneous electron gas of density n. Pinski's model reproduced the qualitative aspects of the Q-V relations, allowing him to suggest that the origin of the Q-V relations lies in a mechanism which exists within the Thomas-Fermi approximation.

2.2.3.2 The Linear Charge Model

Pinski's model was not designed to be quantitatively accurate. Henceforth we will consider models which *were* designed to be quantitatively accurate. The first such model which we will consider is the *linear charge model* (LCM) [28]. This attempts to correctly predict the values of the site charges in *binary alloys*, in which there are only two species of site, which we will refer as A and B. In the LCM, the charge on each site i is given by the following rule:

$$Q_i = 2S_i \lambda_1 N_{i1}, \tag{2.63}$$

where: $S_i = S_A = -1$ if site i belongs to species A; $S_i = S_B = +1$ if site i belongs to species B; N_{i1} is the number of *unlike* nearest neighbours of site i, i.e. if i is an A site then N_{i1} is the the number of nearest neighbour B sites of i and *vice versa* if i is a B site; and λ_1 is a free parameter whose value is not known, and can be determined from *ab initio* methods or experiment. In other words, the charge on a site is linearly proportional to its number of unlike nearest neighbours, with a proportionality constant of $-2\lambda_1$ for A sites and $+2\lambda_1$ for B sites. The LCM was motivated by the reasonable assumption that a site belonging, say, to species A which is surrounded by A sites would have no net charge, as is the case in a pure A metal; it is the presence of unlike neighbours which causes a site to have a non-zero net charge.

The results of order-N calculations have been used both to laud and to criticise the LCM. One of the main criticisms is that the charge of an X site obtained from order-N calculations in random alloys can seemingly be one of a continuous set of values, while in the LCM it can only be one of Z_1 discrete values, where Z_1 is the number of nearest neighbours of any site [23]. This can be seen clearly in Fig. 2.4, where the values of V_i and Q_i for A sites obtained using the LCM (black open circles) are shown. On the other hand, the figure also reveals that the LCM loosely reproduces the linear relationship between V_i and Q_i. In fact, it can be shown analytically that, while the values of V_i and Q_i for sites belonging to the same species do not form an exact Q-V relation in the LCM, the values of $V(Q)_X$ and Q in random alloys do, where $V(Q)_X$ is the average Madelung potential of all X sites with charge Q [29]. Another criticism of the LCM is that the gradients of the Q-V relations for each species, i.e. a_A and a_B, are constrained to take the same value in the LCM, while order-N calculations reveal that there is often a small difference between a_A and a_B [23]. Confusingly, this same result has also been used in support of the LCM: the LCM predicts the same values of a_A and a_B, which is not far from what is seen in the order-N results [29]. A similar situation exists with regards to the fact that order-N calculations reveal that the values of a_A and a_B depend weakly on the concentrations of each species. Proponents of the LCM use this to claim that the λ_1 parameter is independent of concentration - which greatly increases the applicability of the model [29]; while critics of the LCM claim the opposite [23]. In terms of quantitative predictions, it has been shown that, at best, the LCM is only in semiquantitative agreement with the results of order-N calculations. This can be seen by comparing the first two rows of Table 2.2. The first row of the table gives the values for the standard deviation in the charges, and the Q-V relation gradient and intercept for Zn sites in the $Cu_{0.5}Zn_{0.5}$ random alloy obtained from order-N calculations (Ref. [22]). The second row gives the corresponding values predicted by the LCM (Ref. [26]), with the parameter λ_1 calibrated such that the average charge of Zn sites is the same as in the order-N results.

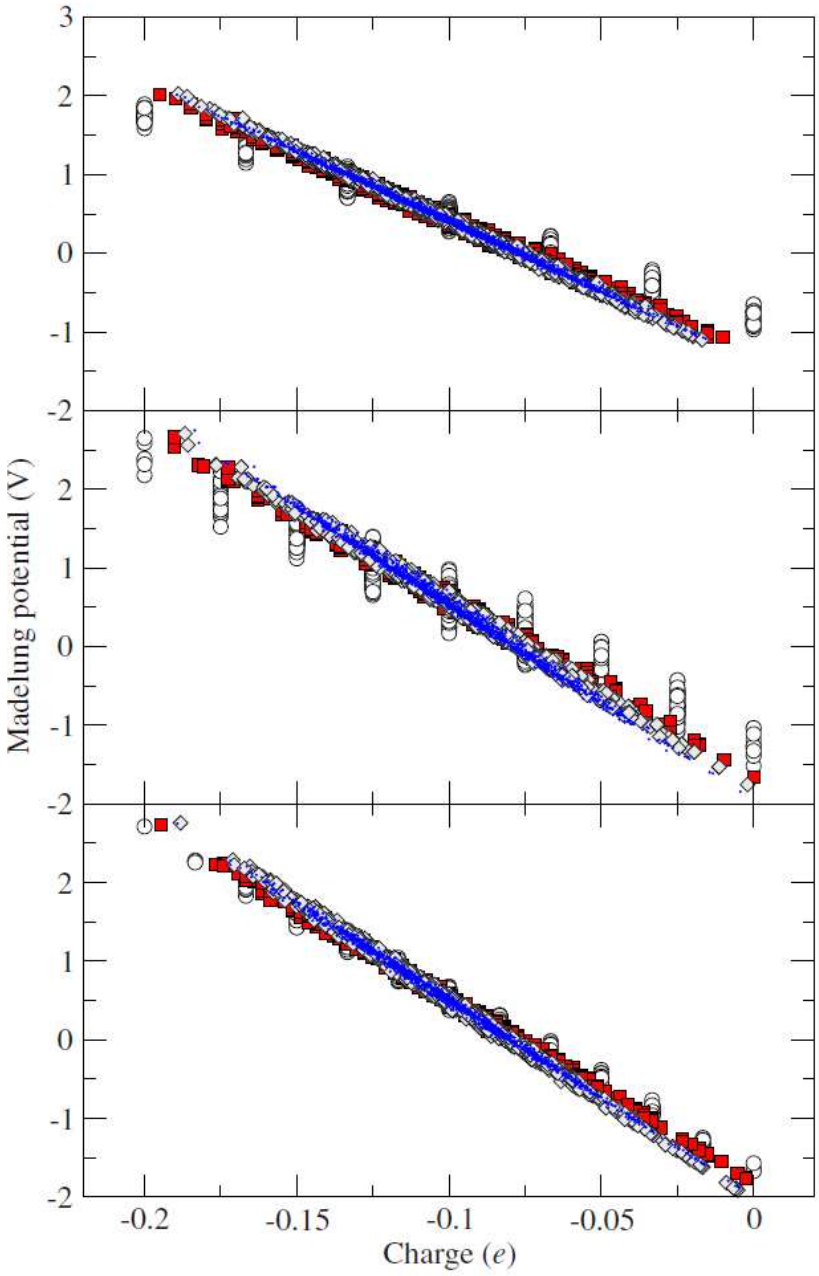


Figure 2.4: V_i vs. Q_i for A sites in sc (top), bcc (middle) and fcc (bottom) random alloys $A_{0.5}B_{0.5}$, whose charges were calculated using the OLCM. The open circles, red squares, light grey diamonds, and blue dots correspond to the OLCM with $\beta_{\max}$ =1, 2, 3 and 4 respectively. Note that the case where $\beta_{\max} = 1$ is exactly equivalent to the conventional LCM. (Taken from Ref. [26]).

	σ	a	k
Order-N calculations	0.026	25.0	-2.00
LCM	0.035	17.3	-1.17
OLCM ($\beta_{\max} = 2$)	0.028	21.7	-1.63
OLCM ($\beta_{\max} = 3$)	0.026	23.5	-1.82
OLCM ($\beta_{\max} = 4$)	0.025	25.0	-1.98
OLCM ($\beta_{\max} = 5$)	0.025	25.0	-1.98

Table 2.2: A comparison between the order-N results for the Zn sites in the $Cu_{0.5}Zn_{0.5}$ random alloy (Ref. [22]), and various models (Ref. [26]). σ, a and k refer to the standard deviation in charges, the negative of the gradient of the Q-V relations, and the intercept of the Q-V relations for the Zn sites. The free parameters in all model results were calibrated such that the average charge of Zn sites is the same as in the order-N results. The values quoted here use units in which charges are measured in units of e, distances are in angstroms, and potentials are in volts.

2.2.3.3 Generalisations of the Linear Charge Model

Despite the shortcomings of the LCM, it certainly performs far better than might be expected from its simplicity, and therefore provides a good starting point for the development of more accurate models. Several generalisations have been proposed. The first of these was the *multi-shell linear charge model* (MLCM) [29], in which the charge on a site is linearly proportional to the number of unlike sites in each of its $\beta_{\max}$ closest nearest neighbour shells, instead of only the first closest nearest neighbour shell:

$$Q_i = 2S_i \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta N_{i\beta}, \tag{2.64}$$

where $N_{i\beta}$ is the number of unlike sites in the βth nearest neighbour shell to site i, and $\lambda_1, \lambda_2, \ldots, \lambda_{\beta_{\max}}$ are free parameters. Setting $\beta_{\max} = 1$ in the MLCM recovers the LCM. In Ref. [29], the free parameters in the model were obtained by fitting the values of Q_i obtained from order-N results to Eqn. (2.64) with $\beta_{\max} = 5$. These values, when used in Eqn. (2.64), were found to give the same relationship between V_i and Q_i as was present in the order-N results. This illustrates that Eqn. (2.64) is capable of reproducing the values of Q_i found in real disordered alloys to a high degree of accuracy. (One could perhaps imagine a relationship between Q_i and the environment of site i other than that described by Eqn. (2.64), but with the same number of adjustable parameters, which does *not* reproduce the real values of Q_i). Despite this, the MLCM has been criticised for the fact that it has too many free parameters, and that it does not address the main conceptual shortcoming of the LCM: the charges of sites belonging to one particular species in the MLCM can take only one of a finite number of values,

while in reality they can seemingly take on any value [23, 30].

In an attempt to reduce the number of free parameters in the MLCM, it was proposed in Ref. [26] that the free parameters be constrained such that the Q-V relations are obeyed 'as closely as is possible' for the chosen value of $\beta_{\max}$.[12] We will refer to this particular case of the MLCM as the *optimised linear charge model* (OLCM). The values of V_i vs. Q_i obtained using the OLCM are shown in Fig. 2.4. As expected, increasing the value of $\beta_{\max}$ causes the (Q_i, V_i) points for each species in the OLCM to increasingly lie upon a single line. In terms of quantitative predictions, the OLCM performs well. As can be seen from Table 2.2, the OLCM, calibrated to give the correct average charge of Zn sites, gives increasingly good quantitative agreement with the order-N results for the random alloy $Cu_{0.5}Zn_{0.5}$ as $\beta_{\max}$ is increased, with no discernible improvement being achieved beyond $\beta_{\max} = 4$. Furthermore, there is a universal mechanism to the screening in the OLCM similar to that described by RS. This is illustrated in Fig. 2.3, where the values of ϕ_β - as defined in Section 2.2.2.2 - for the OLCM are compared to the universal screening curve obtained by RS. It is also clear from the figure that the nature of the screening is only in semi-quantitative agreement with the the results of RS, though interestingly it is in quantitative agreement with the results of Pinski's model. We will reexamine these points in later chapters.

It should be pointed out that other generalisations of the LCM have also been explored in Ref. [26]. However, these proved less fruitful than the OLCM.

2.2.3.4 The Charge-excess Functional Model

An alternative approach to the LCM and its derivatives is the *charge-excess functional model* (CEFM) [31]. Consider the *inter-site* Coulomb energy of the system, i.e. the energy stored in all of the Coulomb interactions *between* sites. This is known as the Madelung energy, which we will denote as E_M. Within the spherical approximation,

$$E_M = \frac{1}{2} \sum_i \sum_j M_{ij} Q_i Q_j. \tag{2.65}$$

The reasoning behind this is along the same lines as our previous discussion regarding Eqn. (2.54). Let E_L denote all contributions to the total energy E of the system *excluding* the Madelung energy. Specifically, E_L consists of the kinetic energy, the exchange-correlation energy, and the *intra-site* Coulomb energy. In the CEFM, E_L is

[12] We will clarify what is meant by 'as closely as is possible', as well as give the specifics of the aforementioned constraint, in the next chapter.

assumed to take the following form:

$$E_L = \frac{1}{2}\sum_i a_i(Q_i - b_i)^2, \tag{2.66}$$

where the parameters a_i and b_i depend only on the chemical species of site i, and take the values a_X and b_X respectively if site i belongs to species X. The fact that the symbols a_i and a_X have already been defined earlier, and that they are gradients of Q-V relations, should be disregarded for now: as we will see in a moment, the apparent clash in notation is logical. An important feature of the above form for E_L - whose origins we will discuss in Chapter 4 - is that the contribution to E_L associated with site i depends only on physical quantities associated with site i itself. Thus the interactions associated with E_L are assumed to be *local* to each site in the CEFM. For this reason we will refer to E_L as the *local energy* and the aforementioned interactions as the *local interactions*. Another important feature of the above equation is that the contribution to E_L associated with site i is quadratic in Q_i, with its minimum at b_i, and its curvature determined by a_i. Hence a_i is a measure of the strength of the local interactions within site i, which act to keep the charge of site i at b_i - the 'bare' charge of the site. The higher a_i is; the stronger these interactions are.

With the above in mind, the values of Q_i in the CEFM are those which minimise the energy function

$$E = E_L + E_M. \tag{2.67}$$

Strictly, we should minimise E subject to the constraint of global charge neutrality, i.e.

$$\sum_i Q_i = 0. \tag{2.68}$$

However, this is unnecessary because, as was shown in Ref. [32], the CEFM energy function automatically yields a charge neutral system when minimised.[13] It was also shown in Ref. [32] that the values of Q_i which minimise E obey Eqn. (2.56) with

$$k_i = a_i b_i. \tag{2.69}$$

In other words, the Q-V relations are implicit in the CEFM: V_i and Q_i for X sites will

[13] In fact, this is only the case for infinite systems [33]. Furthermore, if one does not explicitly impose the constraint of charge neutrality when minimising E, then there is a certain condition which the values of a_i and b_i must meet if the values of Q_i and V_i are to be self-consistent at the minimum in E - as we will see later in Chapter 4.

always form a Q-V relation with gradient a_X and intercept

$$k_X = a_X b_X. \tag{2.70}$$

It is therefore perhaps unsurprising that, if the values of a_X and b_X for each species - which are the free parameters in the CEFM - are derived from order-N calculations, then the CEFM gives an extremely accurate description of disordered alloys [31–33]. We will investigate a particular case of the CEFM in detail in Chapter 4.

2.3 Core Level Shifts in Disordered Alloys

So far we have only discussed variations in local electronic structure in disordered alloys from a theoretical point of view. We will now discuss how they manifest themselves as variations in the *core level shifts* (CLSs) of sites - quantities which can be measured experimentally using XPS. We will explain what a CLS is in a moment, but first it is necessary to understand the basic principles of XPS.

2.3.1 XPS and Photoemission

In an XPS experiment, the system under consideration is irradiated with X-rays of a particular frequency ν. Via individual *photoemission* processes, electrons are ejected from the system. Such electrons are known as *photoelectrons*. The kinetic energies of these photoelectrons are measured, and the distribution of these energies - known as the *XPS spectrum* of the system - reflects the nature of the photoemission processes induced by the X-rays. Photoemission in its entirety is complicated. Here, we will appeal to a simplified picture which is adequate for our needs. For a more accurate description of photoemission, and information on XPS in general, see Ref. [34]. In our simplified picture, we utilise the Hartree-Fock approximation. As mentioned in Section 2.1.1, the Hartree-Fock approximation fails to account for Coulomb correlation. It also fails to account for relativistic effects. However, with regards to CLSs, Coulomb correlation and relativistic effects are inconsequential and can be ignored, as has been pointed out in Ref. [35]. Recall that, in the Hartree-Fock approximation, the electrons can be considered to occupy N_{elec} spin orbitals. For each spin orbital, there is an associated photoemission process. In the process associated with spin orbital p, the system absorbs a photon, resulting in the excitation of an electron from p into a 'free' spin orbital which is not bound to the system. Before photoemission, the system is in what is referred to as the *initial state*, whose total energy we will denote as E^{i}. Immediately after the photoemission, before any of the remaining electrons in the system de-excite to fill the 'hole' left behind in p, the system is in what is referred to as the *final state*, whose total

energy we will denote as E_p^{f}. Energy conservation requires that

$$E_p^{\mathrm{f}} + E_p^{\mathrm{K}} = E^{\mathrm{i}} + h\nu, \tag{2.71}$$

where h is Planck's constant, $h\nu$ is the energy of the photon, and E_p^{K} is the kinetic energy of the photoelectron. The above equation can be rearranged to give

$$E_p^{\mathrm{K}} = h\nu - E_p^{\mathrm{B}}, \tag{2.72}$$

where

$$E_p^{\mathrm{B}} = E_p^{\mathrm{f}} - E^{\mathrm{i}} \tag{2.73}$$

is the *binding energy* associated with p, and is the minimum amount of energy required to eject an electron from the spin orbital. Eqn. (2.72) reveals that, from the distribution of photoelectron kinetic energies obtained from an XPS experiment, one can easily obtain the distribution of the binding energies of the system's spin orbitals. Note that there will be no photoelectrons ejected from spin orbitals for which $E^{\mathrm{B}} > h\nu$, since in this case the photon is unable to provide enough energy for the corresponding photoemission process to occur, and hence the binding energies of such spin orbitals cannot be determined from the experiment.

The above description of photoemission implies that, for a given X-ray frequency, there is a single photoelectron kinetic energy associated with each spin orbital. This is never the case for a number of reasons. Firstly, the final state has a finite lifetime τ; the hole left in p will eventually be filled by another electron in the system, destroying the final state. Due to the energy-time uncertainty principle, it therefore has an uncertainty in its energy $\sim \hbar/\tau$, which, from Eqns. (2.73) and (2.72), implies that there will be an equivalent uncertainty in the photoelectron's kinetic energy. Specifically, the probability of obtaining a photoelectron with a particular kinetic energy is given by a Lorentzian function of width $\sim \hbar/\tau$ centred on the kinetic energy which would be expected by ignoring uncertainty principle considerations.[14] Here, by 'width' of a function we mean its full width at half maximum (FWHM). A second complication is that there are many possible final states associated with photoemission from a single spin orbital.[15] The final state with the lowest energy is that in which the remaining electrons in the sample have *completely relaxed* around the hole such that the energy is minimised. By contrast, in the other possible final states one or more of the following occurs:

- At least one of the remaining electrons in the sample occupies an 'excited' spin

[14] This result can be derived from time-dependent perturbation theory.

[15] In this thesis we always assume that the initial state is the ground state of the system.

orbital, i.e. one which is unoccupied in the aforementioned final state of lowest energy.

- At least one of the remaining electrons in the sample is removed from the sample entirely.[16]
- There are collective oscillations of the remaining electrons, i.e. one or more plasmons is excited.

Photoemission processes associated with these final states always have higher binding energies than that of the 'main' process associated with the final state of lowest energy, and will therefore give rise to photoelectrons with lower kinetic energies than the main process. Those associated with the first and second items in the above list are known as electron shake-up and shake-off processes respectively. In non-metallic systems, individual shake-up and shake-off processes manifest themselves as distinct peaks in the XPS spectrum. This is not the case in metallic systems. Here, there is a continuum of possible unoccupied spin orbitals with energies immediately above the highest-energy occupied spin orbital which can act as 'excited' levels in the sense described above. The result of this is that the binding energies associated with the shake-up and shake-off processes form a continuum above that associated with the main process, which in turn causes the (normally Lorentzian shaped) peak in the XPS spectrum associated with the main process to become asymmetric. The shape of the resulting peak has been deduced by Doniach and Sunjic [36]. The final complication regarding photoemission which must be considered is that the photoelectrons may be scattered inelastically before escaping the system. Inelastic scattering is obviously more likely for photoelectrons created deep within the system, and less likely for photoelectrons created near the system's surface. For this reason, XPS can only probe the system's electronic structure within a finite distance from its surface.

2.3.2 Core Level Shifts

We will now turn to CLSs. Consider a spin orbital which is strongly bound to the system. Such a spin orbital is known as a *core level*; and a *core hole* is a core level which is unoccupied by an electron. Spin orbitals which are not strongly bound to the system are known as *valence levels*. Core levels have the useful property that those bound to an X nucleus in one system are very similar to those bound to an X nucleus in a different system. The reason for this is as follows. An electron occupying a core level p bound to an X nucleus will have a non-zero probability density function $|\psi_p(\mathbf{r}, s)|^2$

[16] An electron being removed from the sample entirely is equivalent to it being excited into a spin orbital which is not bound to the sample.

only for a region located very close to the nucleus, where the one-electron potential is dominated by the nucleus' electrostatic potential. The difference between the *shape* of the one-electron potential in this region and that in the equivalent region for an X nucleus located in a different system will be only very small, due to the fact that the nucleus' electrostatic potentials are so strong in these regions. Therefore the shape of the core levels' probability density functions will be very similar in the two systems. By contrast, the shape of the one-electron potentials experienced by the valence electrons - whose probability density functions extend into the regions between nuclei - will be quite different between systems, and therefore so also will be their probability density functions. Note that the eigenvalues of the core levels in the two systems may be quite different, a point which we will return to later. Furthermore, there is no rigid definition of core and valence electrons: the spin orbitals classified as core levels and those classified as valence levels will depend on context.

The CLS for a core level which is bound to a nucleus belonging to species X is defined as the difference in its binding energy relative to that of a 'reference core level' of the same type which is bound to an X nucleus in a different system. Here, by the 'type' of a core level we are referring to the set of quantum numbers of the atomic core level from which it is derived. E.g. the type of a core level which is derived from the $1s_{1/2}$ atomic core level is $1s_{1/2}$. Note that, while one could work with the absolute binding energies of core levels - which can be determined by an XPS experiment - it is preferable to work with CLSs. This is because CLSs are energy differences, and are expected to be more reliable than absolute binding energies on account of the fact that errors in individual binding energy calculations may cancel when their difference is taken. We will henceforth consider core levels of type t which are bound to X nuclei in an alloy, and will refer to the core level bound to X nucleus i simply as 'core level i'. For such core levels, it is conventional to choose the reference core level to be a t core level bound to a nucleus in a pure X metal. With the above in mind, the CLS of core level i is given by

$$\Delta E_i^{\mathrm{B}} = E_i^{\mathrm{B}} - E_{\mathrm{ref}}^{\mathrm{B}}, \tag{2.74}$$

where E_i^{B} denotes the binding energy of core level i and $E_{\mathrm{ref}}^{\mathrm{B}}$ denotes the binding energy of the reference core level, and an analogous notation will be used for other physical quantities throughout the rest of this section: a particular physical quantity P pertaining to core level i will be denoted as P_i, the analogous property pertaining to the reference core level will be denoted as P_{ref}, and their difference will be denoted as ΔP_i.

2.3.3 Methods for Calculating Core Level Shifts

We will now discuss various methods for determining ΔE_i^{B} theoretically.

2.3.3.1 The Total Energies Method

Substituting Eqn. (2.73) into Eqn. (2.74) gives

$$\Delta E_i^{\mathrm{B}} = (E_i^{\mathrm{f}} - E^{\mathrm{i}}) - (E_{\mathrm{ref}}^{\mathrm{f}} - E_{\mathrm{ref}}^{\mathrm{i}}), \tag{2.75}$$

where E^{i} is the initial state of the alloy, E_i^{f} is the final state of the alloy for core level i, $E_{\mathrm{ref}}^{\mathrm{i}}$ is the initial state of the X metal, and $E_{\mathrm{ref}}^{\mathrm{f}}$ is the final state of the X metal pertaining to the reference core level. Using the above formula, one could determine ΔE_i^{B} from the values of E^{i}, E_i^{f}, $E_{\mathrm{ref}}^{\mathrm{i}}$ and $E_{\mathrm{ref}}^{\mathrm{f}}$ obtained from *ab initio* calculations. This approach, which we will refer to as the *total energies method*, [37] has been widely used to calculate CLSs for disordered alloys in recent years. (See Ref. [38] for a review). The aforementioned calculations utilise the *complete screening picture*, in which it is assumed that the core holes in the final states are completely screened by the valence electrons. In other words, the valence electrons in each final state are fully relaxed around the core hole as to obtain their minimum energy configuration. This is generally expected to be valid given the systems' metallic nature. A common approximation used to simplify such calculations is the *Z+1 approximation* [37]. Defining an *atomic core* as the entity comprising a nucleus and all of its corresponding core levels, in the *Z+1* approximation, a core hole is assumed to act like an extra proton on its corresponding nucleus, and hence an ionised atomic core (i.e. an atomic core with a core hole) with atomic number z is replaced with an unionised atomic core with atomic number $z+1$.[17]

2.3.3.2 The Potential Model

An alternative method to the total energies method, which makes the relationship between CLSs and local electronic structure more transparent, is to use the *potential model*. This model, in various guises, was used successfully in the early days of XPS in order to rationalise CLSs in molecules and compounds, as summarised in Refs. [35,39]. Consider how the alloy under consideration changes as the occupancy η of core level i is varied. Note that at $\eta = 1$, core level i is fully occupied, and the alloy is in the initial state, and that at $\eta = 0$, there is a core hole in core level i, and the alloy is in the final state. The energy E of the alloy is therefore E^{i} at $\eta = 1$, and E_i^{f} at $\eta = 0$. With this in

[17] The *Z+1* approximation is also referred to as the *equivalent core approximation* and the *quasi-atomic model*.

mind, and following Ref. [40], an expansion of E as a Taylor series about $\eta = 1$ yields

$$E_i^{\mathrm{f}} - E^{\mathrm{i}} = -\frac{\partial E}{\partial \eta} + \frac{1}{2}\frac{\partial^2 E}{\partial \eta^2} - \frac{1}{6}\frac{\partial^3 E}{\partial \eta^3} + \ldots, \tag{2.76}$$

where all terms on the right-hand side are evaluated at $\eta = 1$, and similarly for all other physical quantities in this section unless otherwise stated. From Eqn. (2.73), it can be seen that the left-hand side equals E_i^{B}. With this in mind, and applying Janak's theorem (Eqn. (2.48)) to the right-hand side, the above becomes

$$E_i^{\mathrm{B}} = -\varepsilon_i + \frac{1}{2}\frac{\partial \varepsilon_i}{\partial \eta} - \frac{1}{6}\frac{\partial^2 \varepsilon_i}{\partial \eta^2} + \ldots, \tag{2.77}$$

where ε_i is the Kohn-Sham eigenvalue of core level i.[18] Using the same procedure for the reference core level, an analogous result can be derived:

$$E_{\mathrm{ref}}^{\mathrm{B}} = -\varepsilon_{\mathrm{ref}} + \frac{1}{2}\frac{\partial \varepsilon_{\mathrm{ref}}}{\partial \eta} - \frac{1}{6}\frac{\partial^2 \varepsilon_{\mathrm{ref}}}{\partial \eta^2} + \ldots, \tag{2.78}$$

where in this case η pertains to the occupancy of the reference core level. Subtracting the above equation from Eqn. (2.77), we find that

$$\Delta E_i^{\mathrm{B}} = -\Delta\varepsilon_i + \frac{1}{2}\frac{\partial \Delta\varepsilon_i}{\partial \eta} - \frac{1}{6}\frac{\partial^2 \Delta\varepsilon_i}{\partial \eta^2} + \ldots, \tag{2.79}$$

where we have applied Eqn. (2.74) to the left-hand side and exploited the fact that $\partial/\partial\eta$ is a linear operator. Note that the first term on the right-hand side of the above equation depends only on the initial states of the alloy and the X metal: specifically, the initial state (i.e. $\eta = 1$) values of ε_i and $\varepsilon_{\mathrm{ref}}$. By contrast, to evaluate the remaining terms one must know how ε_i and $\varepsilon_{\mathrm{ref}}$ vary with η (at $\eta = 1$) - knowledge which cannot be deduced from the initial states alone. For this reason, the first term is known as the *initial state contribution* to ΔE_i^{B}. We will denote this as $\Delta E_i^{\mathrm{B,i}}$, i.e.

$$\Delta E_i^{\mathrm{B,i}} = -\Delta\varepsilon_i. \tag{2.80}$$

The remaining terms in Eqn. (2.79) collectively form the *final state contribution*, which we will denote as $\Delta E_i^{\mathrm{B,f}}$:

$$\Delta E_i^{\mathrm{B,f}} = \frac{1}{2}\frac{\partial \Delta\varepsilon_i}{\partial \eta} - \frac{1}{6}\frac{\partial^2 \Delta\varepsilon_i}{\partial \eta^2} + \ldots. \tag{2.81}$$

[18] Recall that Janak's theorem pertains not to 'real' electrons, but to quasielectrons. By appealing to this theorem, we are assuming that there is a quasielectron core level which is identical to that of the real core level we are considering, i.e. core level i.

Consider the alloy and the X metal at the same value of η, which is not necessarily $\eta = 1$. In the potential model, we make the following assumptions:

1. The shapes of the wavefunctions of all core levels bound to site i are identical to those of their analogous core levels bound to the reference site.

2. Core level i is localised to a region very close to nucleus i.

Given our earlier discussion regarding core levels in Section 2.3.2, it can be seen that the first of these assumptions is justified, and that the second is justified if core level i is very tightly bound to nucleus i. A consequence of the second assumption is that the one-electron potential 'seen' by core level i can be considered to be a constant U_i, where U_i is the value of the one-electron potential at nucleus i, and similarly for the reference core level. In this case, the two core levels - which, from the first assumption, have the same shape - differ only in their eigenvalues, and the difference in their eigenvalues is equal to the difference in their one-electron potentials. In other words,

$$\Delta\varepsilon_i = \Delta U_i. \tag{2.82}$$

ΔU_i can be split into two contributions: that associated with the exchange-correlation potential, and that associated with the electrostatic potential. The former contribution can be disregarded [35, 40]. The reason for this is as follows. A corollary of the first of the above assumptions is that the core level electron densities within site i in the alloy and the reference site in the X metal are identical.[19] Thus the difference in the 'total' electron density within site i relative to that within the reference site is entirely due to the difference in the valence level electron densities. Now, by definition, valence levels are those spin orbitals which are most weakly bound to the system. This amounts to them occupying the regions between nuclei, and hence the valence level electron densities 'near' nucleus i and the reference nucleus will be small. Since the exchange-correlation potential at $\mathbf{r}$ depends only on the electron density within a region 'near' $\mathbf{r}$,[20] the contribution to the exchange-correlation potentials at the nuclei due to the valence levels will therefore also be small, and we expect the difference in these exchange-correlation potentials to be negligible. Therefore only the contribution to ΔU_i associated with the electrostatic potential need be considered, and hence

$$\Delta U_i = -\Delta V_i^{\mathrm{tot}}, \tag{2.83}$$

where V_i^{tot} denotes the *total* electrostatic potential at nucleus i. The difference in sign

[19] We have tacitly assumed that all core levels bound to a particular nucleus have a vanishing wavefunction outwith the site within which the nucleus is located.

[20] In fact, in the LDA, it depends only on the density at $\mathbf{r}$ itself.

between ΔU_i and $\Delta V_i^{\mathrm{tot}}$ arises because higher electrostatic potentials lower the energy of electrons, and hence the electrostatic contribution to the one-electron potential from a positive charge is negative and *vice versa*. This is reflected in Eqns. (2.1) and (2.34): the external potential due to the positive nuclei is negative, and the Hartree potential due to the negative electrons is positive. We emphasise that the preceding discussion applies for all η, and not just $\eta = 1$. With this in mind, we find that

$$\Delta E_i^{\mathrm{B,i}} = \Delta V_i^{\mathrm{tot}} \tag{2.84}$$

and

$$\Delta E_i^{\mathrm{B,f}} = -\frac{1}{2}\frac{\partial \Delta V_i^{\mathrm{tot}}}{\partial \eta} + \frac{1}{6}\frac{\partial^2 \Delta V_i^{\mathrm{tot}}}{\partial \eta^2} - \dots \tag{2.85}$$

after substituting Eqns. (2.82) and (2.83) into Eqns. (2.80) and (2.81). In the potential model ΔE_i^{B} is therefore understood wholly in terms of the difference in the electrostatic potential at nucleus i relative to that at the reference nucleus.

In this thesis we will focus on the values of $\Delta E_i^{\mathrm{B,i}}$. Consider the initial state of the alloy. V_i^{tot} can be expressed as

$$V_i^{\mathrm{tot}} = \frac{Q_i^{\mathrm{val}}}{r_i^{\mathrm{eff}}} + V_i^{\mathrm{core}} + V_i, \tag{2.86}$$

where Q_i^{val} is the total amount of valence charge (i.e. the charge associated with the valence level electron density) on site i, r_i^{eff} is the *effective radius* at which the charge Q_i^{val} can be considered to reside from nucleus i with regards to the electrostatic potential at the nucleus, and V_i^{core} is the contribution to V_i^{tot} due to the core charge within site i. Recall that V_i is the Madelung potential of site i. Note that the first two terms in the above equation constitute the *intra-site* contribution to V_i^{tot}, while the final term is the *extra-site* contribution. An analogous equation to the above applies for the reference site, which, when subtracted from the above equation, gives

$$\begin{aligned}
\Delta V_i^{\mathrm{tot}} =& \Delta\left(\frac{Q_i^{\mathrm{val}}}{r_i^{\mathrm{eff}}}\right) + \Delta V_i^{\mathrm{core}} + \Delta V_i \\
=& \Delta Q_i^{\mathrm{val}} \frac{1}{r_{\mathrm{ref}}^{\mathrm{eff}}} + Q_{\mathrm{ref}}^{\mathrm{val}} \Delta\left(\frac{1}{r_i^{\mathrm{eff}}}\right) + \Delta Q_i^{\mathrm{val}} \Delta\left(\frac{1}{r_i^{\mathrm{eff}}}\right) + \Delta V_i \\
=& \frac{\Delta Q_i^{\mathrm{val}}}{r_i^{\mathrm{eff}}} + Q_{\mathrm{ref}}^{\mathrm{val}} \Delta\left(\frac{1}{r_i^{\mathrm{eff}}}\right) + \Delta V_i,
\end{aligned} \tag{2.87}$$

where in the second line we have expanded the first term according to Ref. [41] and exploited the fact that $\Delta V_i^{\mathrm{core}} = 0$ on account of the equivalence of the atomic cores

within site i and the reference site, and in the final line we have used the fact that $1/r_i^{\text{eff}} = 1/r_{\text{ref}}^{\text{eff}} + \Delta(1/r_i^{\text{eff}})$. Now, all sites in the X metal are charge neutral in the initial state. There are two consequences of this. Firstly, the Madelung potential of all sites is 0 (due to the spherical approximation) in the X metal, and hence $\Delta V_i = V_i$. Secondly, $\Delta Q_i^{\text{val}} = Q_i$. Therefore

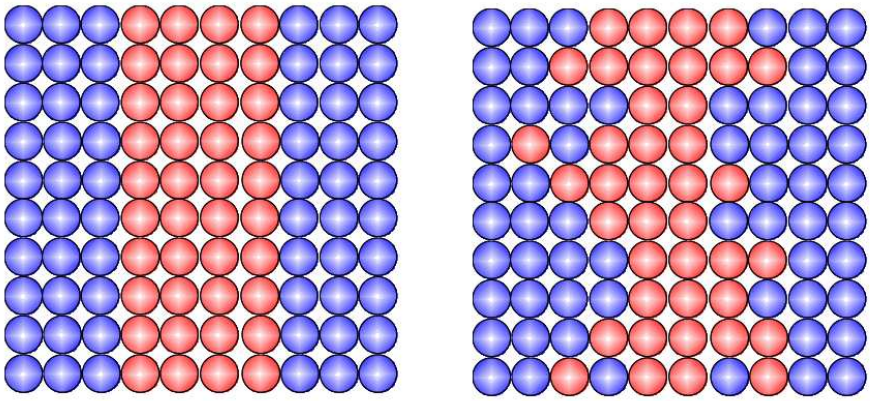
$$\Delta E_i^{\text{B,i}} = \frac{Q_i}{r_i^{\text{eff}}} + V_i + Q_{\text{ref}}^{\text{val}} \Delta \left(\frac{1}{r_i^{\text{eff}}} \right), \tag{2.88}$$

where we have used Eqn. (2.84). Note that the final term is associated with changes in the valence level electron density within site i relative to within the reference site: if $r_i^{\text{eff}} = r_{\text{ref}}^{\text{eff}}$, then the final term is zero; the larger the difference between r_i^{eff} and $r_{\text{ref}}^{\text{eff}}$, the larger the magnitude of the final term. We will use the above equation later in this thesis.

2.3.4 Distribution of CLSs in Disordered Alloys

To conclude this chapter we will discuss the distribution of ΔE_i^{B} for X sites within a particular disordered alloy. While most studies have focused on the mean of this distribution, less attention has been paid to its shape. The latter contains useful information regarding the range of environments in the alloy. The reason for this is as follows. In disordered alloys, the X sites exhibit a variety of environments. Since the CLS for a particular X site depends on its environment, we therefore expect that the X sites will exhibit a variety of CLSs. (The fact that the CLS of a site depends on its environment can be seen most readily through Eqn. (2.88): $\Delta E_i^{\text{B,i}}$ depends on V_i, which is an environment-dependent quantity.) This phenomenon is known as *disorder broadening* on account of the fact that the distribution of CLSs, and hence the distribution of core level binding energies, is broader than for the 'ordered' case of a pure X metal, in which all X sites have the same environment and hence the same core level binding energies. The shape of the CLS distribution therefore reflects the variety of X environments in the alloy. In this thesis, we will focus more on the shape of the CLS distribution than its mean, i.e. we are more interested in the nature of the disorder broadening.

The distribution of CLSs in random alloys has been determined experimentally in Refs. [42–47], and using *ab initio* methods in Refs. [47–49]. More recently, there have also been studies of the distribution of CLSs in more exotic disordered alloys: disordered embedded thin films [38,50]. Ordered embedded thin films consist of several layers of one species embedded in a different host species. In the disordered variety, there is some mixing of atoms at the interfaces between the two species. Schematic



(a) Ordered (b) Disordered

Figure 2.5: Schematic illustrations of an ordered and a disordered embedded thin film.

illustrations of both varieties are illustrated in Fig. 2.5. The studies of random alloys - whose results are summarised in Ref. [27] - have revealed that both initial and final state contributions to the CLSs must be taken into account in order to correctly describe the disorder broadening [49]. This can be seen in Fig. 2.6, where the distributions of Cu $2p_{3/2}$ and Pd $3p_{5/2}$ CLSs in a $Cu_{0.5}Pd_{0.5}$ random alloy are shown along side their initial and final state contributions. Note that the initial and final state disorder broadenings in the Cu core levels reinforce to give a large total disorder broadening; while they partially cancel in the Pd core levels to give a total disorder broadening which is smaller than the initial and final state broadenings individually. Another result to emerge from these studies is that, in some cases, it is necessary to go beyond the assumption which we made at the beginning of Section 2.2 that the positions of the nuclei lie upon an undistorted crystal lattice. This is, as one would expect, only important for alloys in which the atomic sizes of the constituent species exhibit significant size mismatches, such as CuPd and CuAu [47, 49]. Fig. 2.7 shows the CLSs for individual sites in two $Cu_{0.5}Au_{0.5}$ random alloys: one with an undistorted, and one with a distorted crystal lattice. The figure reveals a significant difference in the relationship between a site's CLS and its environment - which is characterised here by its number of unlike nearest neighbours - between the undistorted and distorted cases: lattice distortions act to reverse the (average) relationship between a CLS and its number of unlike nearest neighbours.

When all relevant effects are properly included, the agreement between the results of *ab initio* calculations and experimental results is excellent. However, the reasons *why* the results are what they are is not, in general, very well understood. Models such as those described in Section 2.2.3 and the potential model described in Section 2.3.3.2

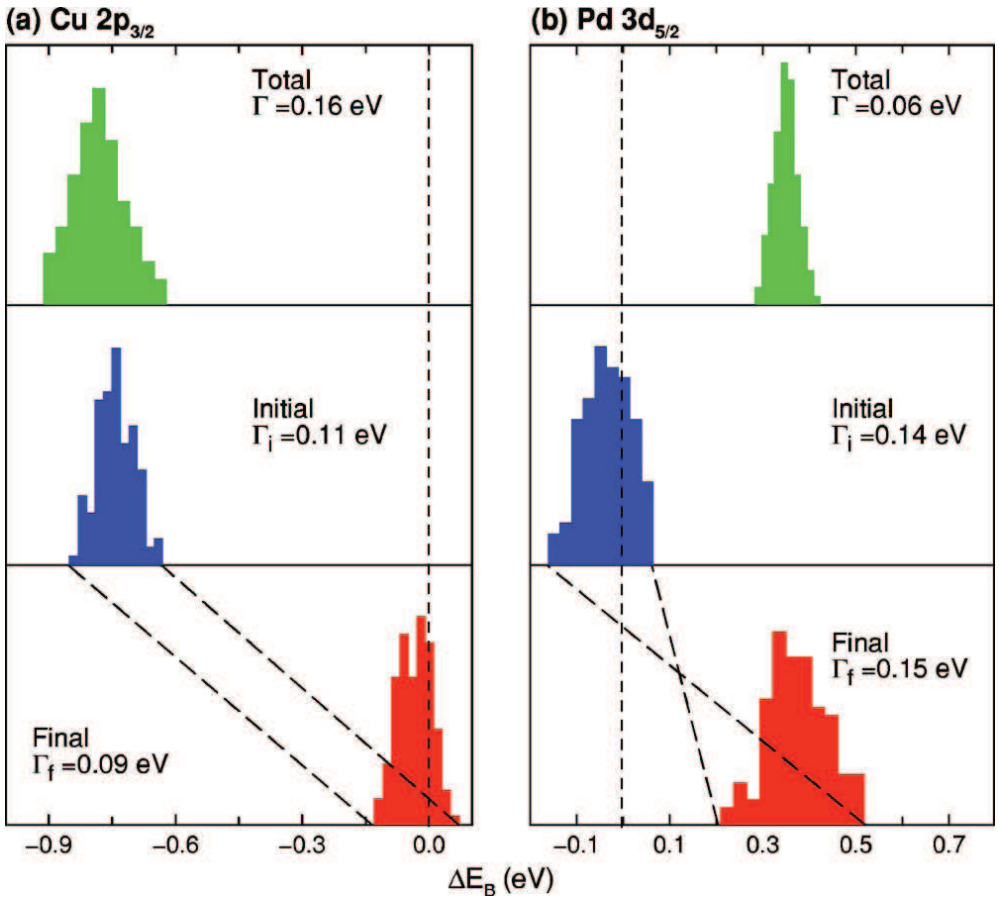


Figure 2.6: The distribution of CLSs (top panel), their initial state contributions (middle panel), and their final state contributions (bottom panel) in a $Cu_{0.5}Pd_{0.5}$ random alloy calculated using the LSGF method. The core level to which each column of panels refers to is indicated above each column. The vales of Γ, Γ_i and Γ_f quoted in each panel refer to the FWHM of the corresponding distribution. The dotted lines connecting the bottom two panels in each column illustrate that: for the Cu $2p_{3/2}$ core levels the sites with the highest and lowest initial state contributions also had the highest and lowest final state contributions respectively; while for the Pd $3p_{5/2}$ core levels the sites with the highest and lowest initial state contributions had the lowest and highest final state contributions respectively. This leads to the initial and final state broadenings to reinforce in the former and partially cancel in the latter. (Taken from Ref. [49].)

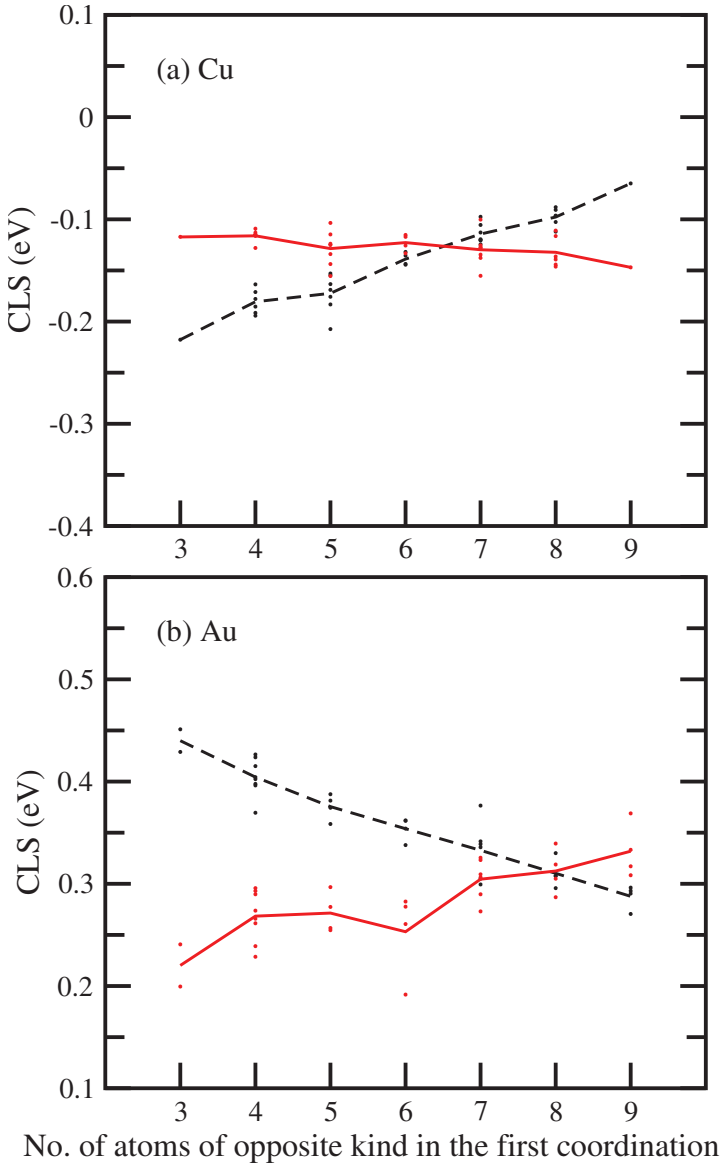


Figure 2.7: The CLSs for sites in the $Cu_{0.5}Au_{0.5}$ random alloy with undistorted (black dashed lines and symbols) and distorted (red solid lines and symbols) crystal lattices, plotted against their number of unlike nearest neighbours. The lines connect the average CLS for each number of unlike neighbours. The CLSs for the Cu $2p_{3/2}$ core levels are shown in (a); and those for the Au $4f_{7/2}$ are shown in (b). These results were obtained using a conventional *ab initio* electronic structure method (i.e. not an order-N method) and a 64 atom cell in which the arrangement of sites belonging to each species was chosen to resemble that of a random alloy as closely as possible. (Taken from Ref. [47].)

have a role to play in making sense of these results. This is especially important given the complexity of the disorder broadening phenomenon: these models provide a means of understanding, individually, the many effects which contribute to the phenomenon as a whole.

Chapter 3

The Optimised Linear Charge Model

The OLCM has previously been used to examine the local electronic structure and distribution of CLSs in the bulk of random alloys. In a random alloy, by construction, the probability of any given site belonging to species X is the same for all sites. In this sense, the concentrations of each species are uniform throughout the entire alloy. In this chapter we will apply the OLCM to various *inhomogeneous* systems. In particular, we will focus on layered systems, in which alloy composition varies along one direction. We will also consider surfaces.

Notation

In this chapter, and throughout the rest of this thesis, we use the following notation unless otherwise stated:

- The set of all sites belonging to species X is denoted simply as X.
- $i \in S$ signifies that site i is in the set of sites S.
- The mean value of some site-dependent property P_i for all sites in the set S is denoted as $\langle P_i \rangle_{i \in S}$.
- The mean value of P_i for all sites in the set S which have charge Q is denoted as $\langle P_i \rangle_{i \in S\,,\, Q_i = Q}$.
- R_β denotes the distance to the βth nearest neighbour shell of any particular site.
- Z_β denotes the number of sites in the βth nearest neighbour shell of any particular site.

- $N_{iX\beta}$ denotes the number of X sites the βth nearest neighbour shell of site i.
- $K_{\beta}^{\delta}(\gamma)$ denotes the number of sites in the δth nearest neighbour shell of any particular site i which are also in the γth nearest neighbour shell of another site j, where sites i and j are in each other's βth nearest neighbour shells.

Furthermore, the βth nearest neighbour shell of any site i will henceforth be referred to simply as 'shell β of i' for the sake of brevity.

3.1 Fundamental Analytical Properties

Before applying the OLCM as described at the beginning of this chapter, it is useful to re-examine the fundamental analytical properties of the model. It should be emphasised that, unless otherwise stated, the forthcoming properties hold for *any* binary alloy, i.e. a binary system with arbitrary site occupations, and not just random alloys.

3.1.1 Screening in the MLCM

We begin by considering the MLCM. As was discussed in Section 2.2.3.3, the OLCM is a particular case of the MLCM, and hence all results - including those presented here - which apply generally to the MLCM will also apply to the OLCM. In the MLCM, the values of Q_i can be obtained by using the following procedure [26], which highlights how charge is screened in the MLCM. Firstly, place an *initial charge* Q_{A0} and Q_{B0} on all A and B sites respectively, where

$$Q_{X0} = 2S_X(1 - c_X) \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta. \tag{3.1}$$

Secondly, distribute a screening charge distribution around each initial charge in the following way: if the initial charge is Q_{X0}, place a charge

$$Q_{X\beta} = -2S_X(1 - c_X)\lambda_\beta \tag{3.2}$$

on each site in shell β of the site containing the initial charge, where $1 \leq \beta \leq \beta_{\max}$.

Proof: We will now show that the procedure described above leads to charges consistent with Eqn. (2.64), which is the expression for Q_i within the MLCM. Consider the charge on an A site i according to the above procedure. The initial charge on i is Q_{A0}. In addition to this, the site obtains Q_{A1} from each A site in shell 1, Q_{B1} from each B site in shell 1, Q_{A2} from each A site in shell 2, Q_{B2} from each B site in shell 2,

etc. Therefore

$$Q_i = Q_{A0} + \sum_{\beta=1}^{\beta_{\max}} (N_{iA\beta} Q_{A\beta} + N_{iB\beta} Q_{B\beta}). \tag{3.3}$$

Substituting Eqns. (3.1) and (3.2) into the above, and simplifying the resulting equation, we find that

$$Q_i = 2 \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta \Big[S_A (1 - c_A)(Z_\beta - N_{iA\beta}) - S_B (1 - c_B) N_{iB\beta} \Big]. \tag{3.4}$$

This becomes

$$Q_i = 2 \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta \Big[S_A (1 - c_A) N_{iB\beta} + S_A c_A N_{iB\beta} \Big]. \tag{3.5}$$

after noting that $S_B = -S_A$, $c_A + c_B = 1$ and $N_{iA\beta} + N_{iB\beta} = Z_\beta$. Simplifying the above equation gives

$$Q_i = 2 S_A \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta N_{iB\beta}. \tag{3.6}$$

Noting that $N_{i\beta} = N_{iB\beta}$ since i is an A site, it can be seen that the above is equivalent to the MLCM expression for Q_i (Eqn. (2.64) with $S_i = S_A$ since i is an A site). This derivation can be repeated with i as a B site to give an equivalent result. ■

Note that the sum of the screening charges for each initial charge Q_{X0} is precisely $-Q_{X0}$, since

$$\sum_{\beta=1}^{\beta_{\max}} Z_\beta Q_{X\beta} = -2 S_X (1 - c_X) \sum_{\beta=1}^{\beta_{\max}} Z_\beta \lambda_\beta = -Q_{X0}, \tag{3.7}$$

where we have used Eqns. (3.1) and (3.2). Therefore the system is necessarily charge neutral in the MLCM.

Let the *localised charge distribution* (LCD) of site i be defined as the charge distribution consisting of site i's initial charge and the associated charge distribution which screens this initial charge. An example of a LCD is illustrated in Fig. 3.1. The procedure described above is therefore equivalent to adding the LCDs associated with each site to the system. Consider now how a perturbation in the charge at site i of size δQ_0 is screened. We can induce such a perturbation by changing the species of site i. Without loss of generality, we will assume that i belongs to species A before the change, and species B after the change. Now, changing the species of i means that the LCD of site i is changed from that of an A site to that of a B site. Therefore

$$\delta Q_0 = Q_{B0} - Q_{A0} \tag{3.8}$$

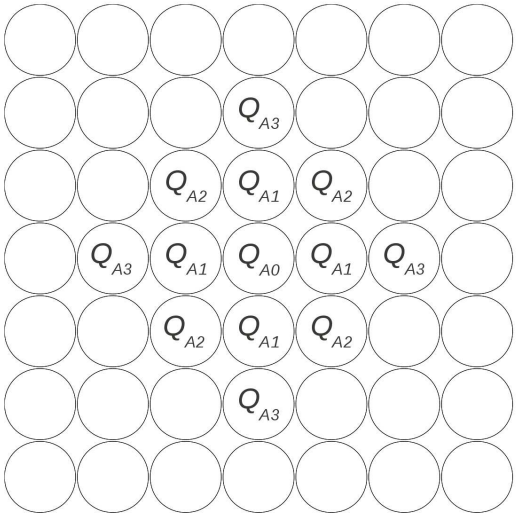


Figure 3.1: An illustration of the localised charge distribution associated with an A site (the central site in the figure) for a square lattice and $\beta_{\max} = 3$. Sites with no specified charge have no charge.

is the change in charge at site i itself, and

$$\delta Q_\beta = Q_{B\beta} - Q_{A\beta} \tag{3.9}$$

is the change in charge for a site in shell $1 \leq \beta \leq \beta_{\max}$ of i. Using Eqns. (3.1) and (3.2), the above equations become

$$\delta Q_0 = -2S_A \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta \tag{3.10}$$

and

$$\delta Q_\beta = 2S_A \lambda_\beta. \tag{3.11}$$

Proof: We will now derive Eqns. (3.10) and (3.11). Substituting Eqns. (3.1) and (3.2) into Eqns. (3.8) and (3.9) gives

$$\delta Q_0 = 2\Big[S_B(1-c_B) - S_A(1-c_A)\Big] \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta \tag{3.12}$$

and

$$\delta Q_\beta = -2\Big[S_B(1-c_B) - S_A(1-c_A)\Big] \lambda_\beta. \tag{3.13}$$

Substituting the relations $S_B = -S_A$ and $1 - c_B = c_A$ into these equations gives Eqns.

(3.10) and (3.11) after simplification. ■

Using Eqns. (3.10) and (3.11), we deduce that the change in charge for a site in shell β, normalised to δQ_0, is

$$\phi_\beta = \frac{\delta Q_\beta}{\delta Q_0} = -\frac{\lambda_\beta}{\sum_{\gamma=1}^{\beta_{\max}} \lambda_\gamma Z_\gamma}. \tag{3.14}$$

Note that ϕ_β is independent of the site i at which the perturbation is located. The screening is therefore the same for all sites within the same system in the MLCM, which is consistent with the *ab initio* findings of RS described in Section 2.2.2.2. We will see later that this is not the case if the system under consideration has a surface - for now, we are assuming that the system extends infinitely in all directions.

Recall that the site charges in the MLCM are postulated to obey Eqn. (2.64). This states that, from each unlike site in shell $1 \leq \beta \leq \beta_{\max}$, site i receives a quantity of charge $2S_i\lambda_\beta$. It is not obvious that the site charges in alloys should be determined by such a mechanism. The above discussion reveals that the MLCM can be reformulated in terms of a single assumption which is more easily justified: that the nature of the screening is the same for all sites. Consider a pure A metal with the same underlying lattice as the alloy we are interested in. From symmetry considerations, it follows that all sites in this metal must have zero net charge. Now, by changing the species of some of the sites to B, one at a time, we can convert the pure A metal into the binary alloy we are interested in. Furthermore, by keeping track of the redistribution of charge after each 'change', we can determine the site charges in the alloy. Note that this procedure applies generally, i.e. it can be used to determine the site charges for *any* model, including the MLCM. Consider now what happens if we change the species of site i. This corresponds to swapping the A atomic core at site i for a B atomic core. Given that the B core will have a different charge than the A core, the change creates an 'impurity' of charge $z_B^* - z_A^*$ within site i, where z_X^* is the total charge of an X core. Note that the impurity's charge depends only on the constituent elements in the system, and hence is the same regardless of where site i comes in the sequence of changes. In response to the impurity, the valence electrons in the system will rearrange themselves, i.e. the valence electrons will *screen* the impurity. Now, given that our system is metallic, to a first approximation, the impurity will be screened in the same way regardless of its location. This is due to the fact that the valence electrons are delocalised, i.e. they are not bound to any particular atomic core, and hence the valence electron gas is well-approximated as being homogeneous.[1] Denoting the change in Q_j induced by the creation of the

[1]Strictly, when changing the species of site i, as well as swapping its A atomic core for a B atomic core, we should also add/remove a few valence electrons to/from the system to keep it charge neutral.

impurity as δQ_0 if $j = i$ and δQ_β if j is in shell $1 \leq \beta \leq \beta_{\max}$ of i, in this case the quantities $\delta Q_0, \delta Q_1, \ldots, \delta Q_{\beta_{\max}}$ are the same regardless of i. This is exactly what was shown earlier to be the case in the MLCM. As alluded to above, if we know how the charge is redistributed after each of the required sites is changed to species B, then we can deduce the site charges for the alloy in question. In the case of the MLCM, this information is contained within the quantities $\delta Q_0, \delta Q_1, \ldots, \delta Q_{\beta_{\max}}$, or equivalently, as can be deduced from Eqn. (3.14), the quantities $\delta Q_0, \phi_1, \phi_2, \ldots, \phi_{\beta_{\max}}$. Thus knowledge of $\delta Q_0, \phi_1, \phi_2, \ldots, \phi_{\beta_{\max}}$, in addition to the site occupations, is sufficient to determine the site charges. In other words, $\delta Q_0, \phi_1, \phi_2, \ldots, \phi_{\beta_{\max}}$ form an alternative set of free parameters for the MLCM to $\lambda_1, \lambda_2, \ldots, \lambda_{\beta_{\max}}$. Note that the two sets are related by Eqns. (3.14) and (3.10).

Before proceeding, it is convenient at this point to mention that the initial charge Q_{X0} defined by Eqn. (3.1) is the mean charge of an X site in the corresponding random alloy, i.e. that with the same underlying crystal structure and concentrations of each species as the system under consideration. We will use this fact later.

Proof: We will now show that Q_{X0} is the mean charge of an X site in the corresponding random alloy. For an A site i, Eqn. (2.64) is

$$Q_i = 2S_A \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta N_{iB\beta}, \tag{3.15}$$

since $N_{i\beta} = N_{iB\beta}$ and $S_i = S_A$. Taking the mean of the above over all A sites gives

$$\langle Q_i \rangle_{i \in A} = 2S_A \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta \langle N_{iB\beta} \rangle_{i \in A}, \tag{3.16}$$

where we have exploited the linearity of the mean: for some set of properties $P_i^1, P_i^2, \ldots, P_i^{m_{\max}}$ pertaining to each site i in a set of sites S whose mean we are interested in, the mean value of the quantity $B + \sum_{m=1}^{m_{\max}} A^m P_i^m$, where $A^1, A^2, \ldots, A^{m_{\max}}$ and B are constants, obeys

$$\left\langle B + \sum_{m=1}^{m_{\max}} A^m P_i^m \right\rangle_{i \in S} = B + \sum_{m=1}^{m_{\max}} A^m \langle P_i^m \rangle_{i \in S}. \tag{3.17}$$

For random alloys

$$\langle N_{iB\beta} \rangle_{i \in A} = Z_\beta c_B \tag{3.18}$$

However, adding/removing a few electrons to the valence electron gas - which consists in its entirety of $\sim 10^{23}$ electrons - has a negligible effect, and hence our failing to do this is inconsequential for any one change of species.

for $\beta \geq 1$. This follows from the fact that, for an A site chosen at random within a random alloy, the probability of it having $N_{iB\beta} = N_{B\beta}$ (for $\beta \geq 1$) is the same as the probability of $N_{B\beta}$ successes in Z_β Bernoulli trials, where each trial has a probability of success of c_B, i.e, the probability $P(N_{B\beta})$ of the site having $N_{iB\beta} = N_{B\beta}$ is determined by the binomial distribution:

$$P(N_{B\beta}) = \binom{Z_\beta}{N_{B\beta}} c_B^{N_{B\beta}} (1 - c_B)^{Z_\beta - N_{B\beta}}. \tag{3.19}$$

The properties of the binomial distribution are such that the mean value of $N_{iB\beta}$ is given by Eqn. (3.18). Substituting Eqn. (3.18) into the above expression for $\langle Q_i \rangle_{i \in A}$ gives

$$\langle Q_i \rangle_{i \in A} = 2S_A(1 - c_A) \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta, \tag{3.20}$$

where we have used the fact that $c_B = 1 - c_A$. From Eqn. (3.1) it can be seen that the right-hand side of the above equation is equal to Q_{A0}. The same arguments lead to the equivalent result for species B. ■

3.1.2 Madelung Potentials in the MLCM

We will now exploit the aforementioned properties of the charge distribution in the MLCM in order to derive expressions for the various contributions to the Madelung potential of any site. Again, since the OLCM is a particular case of the MLCM, what follows also applies to the OLCM. In a moment, the forthcoming expressions will be applied to the OLCM. Consider the contributions to V_i due to each site's LCD. Denoting the contribution to V_i due to the LCD of an X site in shell β of i as $V_{X\beta}$, we find that

$$V_{X\beta} = \begin{cases} -2S_X(1 - c_X) \displaystyle\sum_{\gamma=1}^{\beta_{\max}} \frac{\lambda_\gamma Z_\gamma}{R_\gamma} & \text{if } \beta = 0 \\ 2S_X(1 - c_X) \displaystyle\sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma & \text{if } \beta \geq 1, \end{cases} \tag{3.21}$$

where

$$f_\beta(\gamma) = \frac{Z_\gamma}{R_\beta} - \Sigma_\beta(\gamma), \tag{3.22}$$

$$\Sigma_\beta(\gamma) = \sum_{\delta=1}^{\infty} \frac{K_\beta^\delta(\gamma)}{R_\delta}, \tag{3.23}$$

and we have defined shell 0 of site i as consisting of only site i itself (and hence $R_0 = 0$). Note that the quantities which enter into $f_\beta(\gamma)$ depend only on the underlying lattice, and hence so also will $f_\beta(\gamma)$.

Proof: To derive Eqn. (3.21), we will begin by deriving the expression for V_{A0}. This is the Madelung potential due to the LCD of an A site, at the A site itself. The LCD consists of a charge $Q_{A\gamma}$ being placed on each site in shell $1 \leq \gamma \leq \beta_{\max}$ of the A site. Noting that there are Z_γ sites in shell γ, and that these are at distance R_γ from the A site, it follows that

$$V_{A0} = \sum_{\gamma=1}^{\beta_{\max}} Z_\gamma \frac{Q_{A\gamma}}{R_\gamma}. \tag{3.24}$$

Substituting Eqn. (3.2) into the above gives the expression for V_{A0} in Eqn. (3.21).

We will now derive the expression for $V_{A\beta}$, where $\beta \geq 1$. This is the Madelung potential on any site i due to the LCD of an A site which is in shell $\beta \geq 1$ of i. Consider first the contribution to $V_{A\beta}$ from the charge Q_{A0} on the central site of the LCD. Since the central site is at distance R_β from site i, the contribution to $V_{A\beta}$ is Q_{A0}/R_β. Consider now the contribution to $V_{A\beta}$ from the charges in the LCD of sites in shell $\gamma \geq 1$ of the central site. These each have a charge $Q_{A\gamma}$, though they are not all located at the same distance from site i. Now, since the central site and site i are separated by R_β, $K_\beta^\delta(\gamma)$ is the number of sites in shell γ of the central site which are also in shell δ of i. In fact, this is how $K_\beta^\delta(\gamma)$ is defined. With this in mind, the contribution to $V_{A\beta}$ from sites in shell γ of the central site is

$$\sum_{\delta=1}^{\infty} K_\beta^\delta(\gamma) \frac{Q_{A\gamma}}{R_\delta}. \tag{3.25}$$

By considering all shells $1 \leq \gamma \leq \beta_{\max}$, and adding the contribution from the central site, therefore

$$V_{A\beta} = \frac{Q_{A0}}{R_\beta} + \sum_{\gamma=1}^{\beta_{\max}} \sum_{\delta=1}^{\infty} K_\beta^\delta(\gamma) \frac{Q_{A\gamma}}{R_\delta}. \tag{3.26}$$

Substituting Eqn. (3.2) into the above, and simplifying the resulting equation, we find that

$$V_{A\beta} = 2S_A(1 - c_A) \sum_{\gamma=1}^{\beta_{\max}} \lambda_\gamma \left[\frac{Z_\gamma}{R_\beta} - \sum_{\delta=1}^{\infty} \frac{K_\beta^\delta(\gamma)}{R_\delta} \right]. \tag{3.27}$$

Finally, using Eqns. (3.23) and (3.22), it can be seen that the quantity in the square brackets is $f_\beta(\gamma)$, and hence the above equation is the same as that for $\beta \geq 1$ given in Eqn. (3.21).

This derivation can be repeated using the LCD of a B site in order to derive the

analogous expression for $V_{B\beta}$. ■

Now, V_i can be expressed as a sum of two contributions:

$$V_i = V_i^{\text{local}} + V_i^{\text{distant}}, \tag{3.28}$$

where V_i^{local} is the contribution due to the LCDs of sites within a distance $R_{\beta_{\max}}$ from site i (including site i itself), and V_i^{distant} is the contribution due to the LCDs of sites beyond distance $R_{\beta_{\max}}$ from site i. We will refer to the regions within and outwith $R_{\beta_{\max}}$ of site i as the *local environment* and *distant environment* of site i respectively; and hence will refer to V_i^{local} as the *local Madelung potential* of site i and V_i^{distant} as the *distant Madelung potential* of site i. We emphasise that the local Madelung potential of site i is *not* the contribution to V_i due to the charge on sites within distance $R_{\beta_{\max}}$ of site i, but is the Madelung potential due to the LCDs of sites within this distance. Similar applies for the distant Madelung potential. Explicitly, the two contributions are:

$$V_i^{\text{local}} = V_{A0} + \sum_{\beta=1}^{\beta_{\max}} (V_{A\beta} N_{iA\beta} + V_{B\beta} N_{iB\beta}) \tag{3.29}$$

and

$$V_i^{\text{distant}} = \sum_{\beta=\beta_{\max}+1}^{\infty} (V_{A\beta} N_{iA\beta} + V_{B\beta} N_{iB\beta}), \tag{3.30}$$

where without loss of generality we have assumed that i is an A site. The equivalent expressions derived below for a B site can be obtained by interchanging all A indices for B indices and *vice versa*. Using Eqn. (3.21), it can be shown that Eqn. (3.29) becomes

$$V_i^{\text{local}} = -2S_A \sum_{\beta=1}^{\beta_{\max}} \sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma) \lambda_\gamma N_{i\beta} + 2S_A(1-c_A) \sum_{\beta=1}^{\beta_{\max}} Z_\beta \sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma) \lambda_\gamma + V_{A0}. \tag{3.31}$$

We will use this expression in a moment.

Proof: We will now derive Eqn. (3.31). Consider the quantity in parenthesis in Eqn. (3.29). Using Eqn. (3.21), we find that

$$V_{A\beta} N_{iA\beta} + V_{B\beta} N_{iB\beta} = 2 \sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma) \lambda_\gamma \Big[S_A(1-c_A) N_{iA\beta} + S_B(1-c_B) N_{iB\beta} \Big]. \tag{3.32}$$

Using the relations $S_B = -S_A$, $N_{iA\beta} + N_{iB\beta} = Z_\beta$ and $c_A + c_B = 1$, this becomes

$$V_{A\beta}N_{iA\beta} + V_{B\beta}N_{iB\beta} = 2\sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma\Big[S_A(1-c_A)(Z_\beta - N_{iB\beta}) - S_A c_A N_{iB\beta}\Big], \quad (3.33)$$

which in turn becomes

$$V_{A\beta}N_{iA\beta} + V_{B\beta}N_{iB\beta} = 2S_A\sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma\Big[(1-c_A)Z_\beta - N_{iB\beta}\Big] \quad (3.34)$$

after simplification. Substituting the above into Eqn. (3.29) gives

$$V_i^{\text{local}} = V_{A0} + 2S_A\sum_{\beta=1}^{\beta_{\max}}\sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma\Big[(1-c_A)Z_\beta - N_{iB\beta}\Big], \quad (3.35)$$

which, after expanding the square brackets, gives Eqn. (3.31). ■

3.1.3 The OLCM Constraint

We now turn to the OLCM. In Section 2.2.3.3 it was mentioned that the OLCM is the particular case of the MLCM in which the free parameters are constrained such that the Q-V relations are obeyed 'as closely as is possible' for the chosen value of $\beta_{\max}$. However, we did not clarify what was meant by 'as closely as is possible', nor did we specify what this constraint is. The constraint is as follows: in the OLCM, the vector

$$\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_{\beta_{\max}}) \quad (3.36)$$

- which comprises all of the MLCM free parameters - must be an eigenvector of the $\beta_{\max} \times \beta_{\max}$ 'f-matrix' whose (β,γ)th element is $f_\beta(\gamma)$, where recall that $f_\beta(\gamma)$ is a lattice-dependent quantity defined by Eqns. (3.22) and (3.23). In other words,

$$\sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma = a\lambda_\beta \quad \text{for } 1 \leq \beta \leq \beta_{\max}, \quad (3.37)$$

where we have denoted the eigenvalue of $\boldsymbol{\lambda}$ as a. Recall that the motivation behind this constraint was to reduce the parameter space of the MLCM. This is achieved as follows: in the general MLCM $\boldsymbol{\lambda}$ can be any vector in the $\beta_{\max}$-dimensional Euclidean vector space $\mathbb{R}^{\beta_{\max}}$; while in the OLCM $\boldsymbol{\lambda}$ can be any vector in the eigenspace of the $\beta_{\max} \times \beta_{\max}$ f-matrix, which itself is a subspace of $\mathbb{R}^{\beta_{\max}}$.

3.1.4 Madelung Potentials in the OLCM

We will now show why the aforementioned constraint leads to the Q-V relations being obeyed 'as closely as is possible' for the chosen value of $\beta_{\max}$. Substituting the above equation into Eqn. (3.31) gives

$$V_i^{\text{local}} = -2S_A a \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta N_{i\beta} + 2S_A(1-c_A)a \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta + V_{A0}. \tag{3.38}$$

From Eqns. (2.64) and (3.1), this can be seen to be equivalent to

$$V_i^{\text{local}} = -aQ_i + k_A, \tag{3.39}$$

where

$$k_A = aQ_{A0} + V_{A0}. \tag{3.40}$$

Alternatively, using Eqns. (3.1) and (3.21), k_A can be expressed as

$$k_A = 2S_A(1-c_A)\left[a \sum_{\gamma=1}^{\beta_{\max}} \lambda_\gamma Z_\gamma - \sum_{\gamma=1}^{\beta_{\max}} \frac{\lambda_\gamma Z_\gamma}{R_\gamma}\right]. \tag{3.41}$$

Eqn. (3.39) reveals that the OLCM constraint on $\boldsymbol{\lambda}$ leads to the local Madelung potentials V_i^{local} obeying Q-V relations. However, the 'total' Madelung potentials do not obey Q-V relations, due to the fact that the distant Madelung potentials V_i^{distant} will vary between sites belonging to the same species with the same charge. This is made clear by the expression

$$V_i = -aQ_i + k_A + V_i^{\text{distant}}, \tag{3.42}$$

which follows from Eqn. (3.28) and (3.39). The reason for this is that the charge of any site depends only on the values of $N_{i\beta}$ for those shells β in its local environment, and not on shells in its distant environment. With regards to reproducing the Q-V relations, this is, however, the best that we can do given the constraint that the charge of any site depends only on the occupation of shells within a finite distance. This notwithstanding, it may be that in practice the magnitudes of V_i^{distant} are negligible, in which case the Q-V relations can be considered to hold for all practical purposes. This occurs if the magnitudes of λ_β 'converge' to zero by $\beta = \beta_{\max}$. The reasoning behind this is as follows. Consider the LCD of an X site. As described earlier, $V_{X\beta}$ is the Madelung potential of site i due to the LCD of an X site in shell β. Equivalently, $V_{X\beta}$ is the Madelung potential of a site at R_β from the origin due to the LCD of an X

site at the origin. Consider $V_{X\beta}$ for shells β which are increasingly far away from the origin. Applying Eqn. (3.37) to Eqn. (3.21), we find that

$$V_{X\beta} = 2S_X(1 - c_X)a\lambda_\beta \quad \text{if } 1 \leq \beta \leq \beta_{\max}. \tag{3.43}$$

From this we see that if λ_β converges to zero by $\beta = \beta_{\max}$, then so also must $V_{X\beta}$. If this is the case then the LCD is such that the Madelung potential due to the charge Q_{X0} on the site at the origin is perfectly screened at a distance $R_{\beta_{\max}}$ by the remaining site charges in the LCD, and it follows that $V_{X\beta}$ will also be zero for $\beta > \beta_{\max}$. Since the distant potential of any particular site is a linear combination of $V_{A\beta}$ and $V_{B\beta}$ for $\beta > \beta_{\max}$ (see Eqn. (3.30)), then this in turn implies that the distant potential will vanish for all sites if λ_β converges to zero by $\beta = \beta_{\max}$.

While the Q-V relations do not hold exactly, they do hold 'on average' if we consider all A sites with the same charge Q, so long as the mean values of $N_{i\beta}$ for such sites obey

$$\langle N_{i\beta}\rangle_{i\in A\,,\,Q_i=Q} = Z_\beta c_B \quad \text{for all } \beta > \beta_{\max}; \tag{3.44}$$

in other words, if the average composition of each shell in the distant environment of the sites reflects the composition of the whole system. The above equation holds in random alloys, and in alloys with short-range order - so long as the range of the short-range order is less than $R_{\beta_{\max}}$. The reason for this is along the same lines as the discussion immediately following Eqn. (3.18): in such systems, the values of $N_{i\beta}$ for $\beta > \beta_{\max}$ of all A sites with charge Q are weighted according to the binomial distribution. Using the above equation, we find that the mean Madelung potential of the aforementioned set of sites is

$$\langle V_i\rangle_{i\in A\,,\,Q_i=Q} = -aQ + k_A, \tag{3.45}$$

which reveals that, for each possible charge Q that A sites could have, the values of $\langle V_i\rangle_{i\in A\,,\,Q_i=Q}$ and Q obey a Q-V relation.

Proof: The derivation of Eqn. (3.45) is as follows. Taking the mean of Eqn. (3.42) over all A sites with charge Q gives

$$\langle V_i\rangle_{i\in A\,,\,Q_i=Q} = -a\langle Q_i\rangle_{i\in A\,,\,Q_i=Q} + k_A + \langle V_i^{\text{distant}}\rangle_{i\in A\,,\,Q_i=Q}, \tag{3.46}$$

where we have exploited the linearity of the mean (Eqn. (3.17)). Noting that $\langle Q_i\rangle_{i\in A\,,\,Q_i=Q} = Q$ gives

$$\langle V_i\rangle_{i\in A\,,\,Q_i=Q} = -aQ + k_A + \langle V_i^{\text{distant}}\rangle_{i\in A\,,\,Q_i=Q}. \tag{3.47}$$

Comparing the above equation to Eqn. (3.45), we see that what remains to be shown is that

$$\langle V_i^{\text{distant}} \rangle_{i \in A\,,\, Q_i = Q} = 0. \tag{3.48}$$

Consider Eqn. (3.30). Using Eqn. (3.34), this becomes

$$V_i^{\text{distant}} = 2S_A \sum_{\beta=\beta_{\max}+1}^{\infty} \sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma \Big[(1-c_A)Z_\beta - N_{iB\beta}\Big]. \tag{3.49}$$

Taking the mean of this over all A sites with charge Q gives

$$\langle V_i^{\text{distant}} \rangle_{i \in A\,,\, Q_i = Q} = 2S_A \sum_{\beta=\beta_{\max}+1}^{\infty} \sum_{\gamma=1}^{\beta_{\max}} f_\beta(\gamma)\lambda_\gamma \Big[(1-c_A)Z_\beta - \langle N_{i\beta} \rangle_{i \in A, Q}\Big], \tag{3.50}$$

where we have again exploited the linearity of the mean, and also used the fact that $N_{iB\beta} = N_{i\beta}$ for A sites. Substituting Eqn. (3.44) into the quantity in the square brackets and noting that $c_B = 1 - c_A$, we see that it vanishes for all $\beta > \beta_{\max}$, and hence Eqn. (3.48) holds. ■

3.2 Details of Computational Calculations

Having described the fundamental analytical properties of the model, in the remainder of this chapter we will discuss results of new computational calculations utilising the OLCM for a variety of systems. In this section we give the details of the calculations. Following the convention used in experimental studies, we will often quote charges in units of e, distances in angstroms, potentials in volts and energies in eV. The last three rows in Table 2.1 can be used to convert between quantities utilising these units and their corresponding quantities in HAU.

All systems considered in the calculations had a fcc underlying lattice with $R_1 = 3$Å, and were characterised by their concentration profiles in the z-direction. In other words, each system was characterised by the concentrations of A sites in each of its 001 planes, which we will refer to as *layers*. We will denote the concentration of A sites in layer l as c_A^l. The systems were all periodic, with the repeating unit consisting of N_x, N_y and N_z conventional fcc unit cells in the x, y and z-directions respectively. Thus the repeating unit is a cuboid with x, y and z dimensions N_xD, N_yD and N_zD respectively, where $D = \sqrt{2}R_1$ is the lattice parameter of the conventional fcc unit cell. Within each of the $2N_z$ layers - which are separated by distance $d = D/2$ - the species of the $N_{\text{layer}} = 2N_xN_y$ sites in each layer l were assigned at random such that the total number of A and B sites reflected the layer concentrations c_A^l and c_B^l. The systems

Quantity	Value
λ_1	0.007190 e
λ_2	0.001413 e
λ_3	2.272×10^{-4} e
λ_4	-1.776×10^{-5} e
a	21.44 V/e
$\sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta$	0.1 e
$\sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta / R_\beta$	0.4575 V

Table 3.1: The elements of $\boldsymbol{\lambda}$ which were used in the computational calculations and related quantities. The determination of $\boldsymbol{\lambda}$ is described in the text.

are therefore approximations of disordered alloys in which each layer l is itself a 2D random alloy with a square underlying lattice, nearest neighbour distance of 3 Å, and concentration of A sites c_A^l. The accuracy of the approximation is higher, the larger the size of the repeating unit. Most systems considered had $N_z = \infty$, i.e. the unit cell was infinite in extent along the z-axis. Such systems are therefore periodic only in 2 dimensions; specifically, in the xy-plane.

The value of $\beta_{\max}$ used in the calculations was 4, and the vector $\boldsymbol{\lambda}$ was chosen to be the eigenvector of the underlying lattice's 4×4 f-matrix which had the highest eigenvalue a, with a normalisation such that $Q_{A0} = -0.1e$ at $c_A = 0.5$.[2] The elements of this vector and related quantities are given in Table 3.1. Note that the aforementioned choices of R_1 and Q_{A0} are merely 'reasonable' choices which do not necessarily pertain to any specific alloy.[3] The same applies to the parameters introduced in a moment regarding the simulation of XPS spectra. This is justified because the goal of the calculations was not to yield quantitative predictions for particular systems, but to gain an appreciation of the phenomena exhibited in a wide range of systems. In the calculations, charges were allocated to each site according to Eqn. (2.64). In the case of systems with surfaces, and hence a vacuum region containing no sites, 'missing' sites in the vacuum region were not counted as unlike sites for any species, i.e. the vacuum region was ignored in the allocation of site charges in such systems. Furthermore, we have assumed a 'clean' termination of the crystal lattice at the surface, i.e. we ignore phenomena such as surface reconstruction or relaxation of the surface layer towards the bulk of the alloy.

We will now discuss how the Madelung potentials were calculated. It is well known

[2]The fact that there is flexibility in the choice of eigenvalue will be discussed in detail at the end of this chapter.

[3]Our notion of 'reasonable' with regards to Q_{A0} was formulated with Ref. [23] in mind. In this study, $Cu_{0.5}Pd_{0.5}$ and $Cu_{0.5}Zn_{0.5}$ random alloys were found - through the use of *ab initio* calculations - to have $Q_{A0} \approx -0.1e$.

that, for periodic systems, the evaluation of V_i as defined by Eqns. (2.54) and (2.55) is problematic. Specifically, the summation over j in Eqn. (2.54) is conditionally convergent, which means that one can obtain any value for the summation by altering the order in which its various terms are added together. This is a technical problem caused by the long-range nature of the electrostatic interactions. (See Ref. [4] for more details.) The Ewald summation method [51] addresses this problem. The physical picture behind this method is as follows. Firstly, to each charge Q_j one adds a screening charge distribution which is Gaussian in shape, centred on $\mathbf{R}_j$, and of standard deviation $1/(\sqrt{2}\eta)$.[4] The result of this is to make the electrostatic interactions short-ranged in real space, i.e. in order to evaluate the Madelung potential of any particular site i, we only need consider the contribution from the charges Q_j and their corresponding screening charge distributions within a finite distance from i. However, to get the 'true' Madelung potential for site i we must subtract the electrostatic potential at $\mathbf{R}_i$ due to the Gaussian screening charge distributions. The electrostatic potential due to these distributions, while long-ranged in real space, is short-ranged in reciprocal space, for which reason it is evaluated in reciprocal space, before being subtracted from the result of the aforementioned real space calculation to get the desired result. The value of η determines the relative ranges of the real and reciprocal space interactions. The larger η is, the better the screening is, and hence the shorter the range of the interactions in the real space calculation. Contrariwise, the larger η is, the less smooth the screening charge distributions are, and hence the longer the range of the interactions associated with these distributions in reciprocal space. The opposite is true for smaller values of η. Hence if η is large then one need only include a few image repeating units in the real space calculation, but must include many in the reciprocal space calculation, and *vice versa* if η is small. For all systems we considered, for the chosen value of η, the extent of the real and reciprocal space summations was always chosen such that all Madelung potentials in the system were converged to within 1×10^{-6} V. The relevant expression for V_i within the Ewald summation method for a system with 3D periodicity is [52]

$$V_i = \sum_j Q_j \Bigg\{ \sum_{\mathbf{T}}{}' \frac{\operatorname{erfc}\big(\eta|\mathbf{R}_{ij}+\mathbf{T}|\big)}{|\mathbf{R}_{ij}+\mathbf{T}|} - \frac{2\eta}{\sqrt{\pi}}\delta_{ij} - \frac{\pi}{\eta^2\Omega} + \frac{4\pi}{\Omega}\sum_{\mathbf{G}\neq 0}\frac{1}{|\mathbf{G}|^2}\exp\left(-\frac{|\mathbf{G}|^2}{4\eta^2}\right)\cos(\mathbf{R}_{ij}\cdot\mathbf{G})\Bigg\}, \tag{3.51}$$

where: $\mathbf{R}_{ij} = \mathbf{R}_i - \mathbf{R}_j$, the summation over j is over all sites in the same repeating

[4] Note that, as opposed to in Chapter 2 and Chapter 5, η does *not* refer to the occupancy of a particular spin orbital in this chapter.

unit as i, the set of vectors $\{\mathbf{T}\}$ is the set of all possible translational vectors taking a point in one repeating unit to the corresponding point in an image repeating unit, the prime on the summation over all $\mathbf{T}$ indicates that the $\mathbf{T} = \mathbf{0}$ term is ignored if $j = i$, the set of vectors $\{\mathbf{G}\}$ are the reciprocal lattice vectors corresponding to $\{\mathbf{T}\}$, Ω is the volume of the repeating unit, and $\mathrm{erfc}(x)$ is the complementary error function. Similarly, for systems in which the repeating unit is a cuboid of dimension L in the x and y directions, and infinitely along the z-axis, the expression for V_i is [53]

$$\begin{aligned} V_i = \sum_j Q_j \Bigg\{ & \sum_{\mathbf{T}}{}' \frac{\mathrm{erfc}\big(\eta|\mathbf{R}_{ij}+\mathbf{T}|\big)}{|\mathbf{R}_{ij}+\mathbf{T}|} - \frac{2\eta}{\sqrt{\pi}}\delta_{ij} - \frac{2\pi}{L^2} z_{ij}\,\mathrm{erf}(\eta z_{ij}) \\ & + \frac{2\pi}{L^2} \sum_{\mathbf{G}\neq 0} \frac{1}{|\mathbf{G}|} \exp\big(|\mathbf{G}|z_{ij}\big)\,\mathrm{erfc}\left(\frac{|\mathbf{G}|}{2\eta} + \eta z_{ij}\right) \cos(\boldsymbol{\rho}_{ij}\cdot\mathbf{G}) \\ & - \frac{2\sqrt{\pi}}{L^2\eta} \exp\Big[-(\eta z_{ij})^2\Big] \Bigg\}, \end{aligned} \tag{3.52}$$

where: $\boldsymbol{\rho_i}$ denotes the position of site i in the xy plane, z_i denotes the z coordinate of site i, $\boldsymbol{\rho}_{ij} = \boldsymbol{\rho}_i - \boldsymbol{\rho}_j$, $z_{ij} = z_i - z_j$, $\mathrm{erf}(x)$ is the error function, the set of vectors $\{\mathbf{T}\}$ is in this case the set of lattice vectors for a 2D square lattice of dimension L, and the set of vectors $\{\mathbf{G}\}$ is the corresponding set of 2D reciprocal lattice vectors.

From the values of V_i and Q_i determined as described above, we have also used the potential model (see Section 2.3.3.2) to calculate the *initial state CLSs* associated with each site. We will henceforth, for the sake of brevity, use the terms 'initial state CLS' and 'CLS' interchangeably with regards to our results. However, the term 'CLS' used in any other context, i.e. when referring to *ab initio* or experimental results, should be taken to mean a 'true' CLS including both initial and final state contributions. The relevant equation is Eqn. (2.88). Unfortunately, while the OLCM provides Q_i and V_i, it does not provide r_i^{eff}, which is required in order to calculate the first and final terms in the aforementioned equation. Following Refs. [42, 43, 54], we will assume that r_i^{eff} takes the plausible value of $R_1/2$ for all sites. This assumption enables us to calculate the first term in Eqn. (2.88), and reduces the final term to a species-dependent constant, which, for species X, is

$$\Theta_X = Q_{\mathrm{ref},X}^{\mathrm{val}} \left(\frac{2}{R_1} - \frac{1}{r_{\mathrm{ref},X}^{\mathrm{eff}}} \right), \tag{3.53}$$

where $Q_{\mathrm{ref},X}^{\mathrm{val}}$ and $r_{\mathrm{ref},X}^{\mathrm{eff}}$ are the amount of valence charge on any site in, and the effective distance of, a pure X metal respectively. Note that Θ_X depends on how the shape of the valence level electron density differs within an X site in the alloy under consideration relative to within a site in a pure X metal. In our calculations we have ignored the

contribution to the CLSs due to Θ_X, and hence used the formula

$$\Delta E_i^{\mathrm{B,i}} = \frac{2Q_i}{R_1} + V_i \tag{3.54}$$

for all sites. Thus our CLSs are, for X sites, measured relative to Θ_X; or equivalently, we have set the 'zero' of the CLS scale for X sites to Θ_X. Since we are interested in the *range* of CLSs, and not their absolute values, this is unimportant. With CLSs determined using the above equation, we have simulated the associated spectrum which would be observed in an XPS study. This was done using the following formula:

$$I_X(E) = G(E;\Gamma_{\mathrm{E}}) \otimes \left[\sum_{i\in X}{}' e^{-z_i/\ell} L(E - \Delta E_i^{\mathrm{B,i}};\Gamma_{\mathrm{L}}) \right], \tag{3.55}$$

where $I_X(E)$ is the intensity of the spectrum corresponding to species X at CLS E, $G(E;\Gamma_{\mathrm{E}})$ denotes a Gaussian function with FWHM Γ_{E}, $L(E;\Gamma_{\mathrm{L}})$ denotes a Lorentzian function with FWHM Γ_{L}, the symbol $\otimes$ denotes the convolution operator, Γ_{E} is the FWHM associated with instrumental broadening, Γ_{L} is the FWHM associated with lifetime broadening of the core level under consideration, z_i is the distance of site i from the surface,[5] ℓ is the inelastic mean free path of the photoelectrons, and the prime on the summation signifies that sites in the surface layer, i.e. the layer of sites immediately adjacent to the vacuum region, are to be ignored. It is well known that in pure metals the CLSs of core levels bound to a site in the surface layer are different to the corresponding core levels bound to a site within the bulk of the metal. The difference in these quantities is known as the *surface core level shift*, and is ~ 0.5 eV in magnitude, with a positive or negative sign [55]. It is to avoid this complication, which also occurs in alloys, that we ignore the contribution from sites in the surface layer in Eqn. (3.55). Eqn. (3.55) therefore gives the 'bulk' spectrum of the system, where by 'bulk' in this context we mean 'not the surface layer'. The factor $e^{-z_i/\ell}$ in Eqn. (3.55) accounts for the fact that, on account of inelastic scattering, photoelectrons originating from sites at different distances from the surface will have different probabilities of escaping the alloy and being detected. For all simulated spectra we used $\Gamma_{\mathrm{L}} = 0.3$ eV and $\Gamma_{\mathrm{E}} = 0.3$ eV. For systems in which surfaces are explicitly included in the system, which we will refer to as *surface systems*, we used $\ell = 10$ Å; for systems in which surfaces are not explicitly included in the system, i.e. systems which by construction extend infinitely in all directions, which we will refer to as *bulk systems*, we used $\ell = \infty$.

[5] Note that in this chapter z_i does not refer to the atomic number of nucleus i, as it does in other chapters of this thesis.

In the latter case, Eqn. (3.55) simplifies to

$$I_X(E) = G(E;\Gamma_{\mathrm{E}}) \otimes \left[\sum_{i \in X} L(E - \Delta E_i^{\mathrm{B,i}};\Gamma_{\mathrm{L}}) \right]. \tag{3.56}$$

Note that our simulated spectra do not account for all of the features which are present in real XPS spectra, such as Doniach-Sunjic asymmetry. Since each simulated spectrum serves only to illustrate how differences in environments affect the distribution of initial state CLSs, we are justified in ignoring such complications.

3.3 Bulk Systems

3.3.1 Random Alloys

We will begin by considering the bulk of random alloys. Plots of V_i vs. Q_i for A sites are given in Fig. 3.2 for random alloys with various values of c_A. The repeating unit in these calculations had $N_x = N_y = N_z = 5$, which corresponds to 500 sites, and $c_A^l = c_A$ for all layers. For comparison, the analytical predictions for $\langle V_i \rangle_{i \in A\,, Q_i = Q}$ vs. Q of Eqn. (3.45) are also shown in the figure, as well as a line which connects the points (Q_{A0}, V_{A0}) for each concentration. Recall that, for a given value of c_A, Q_{A0} is the mean charge of an A site in a random alloy. It turns out that V_{A0} is the mean Madelung potential of an A site, and hence the aforementioned line connects the analytically determined mean (Q_i, V_i) points for each concentration.

Proof: We will now show that V_{A0} is the mean Madelung potential of an A site in a random alloy. Consider Eqn. (3.42). Taking the mean of this over all A sites gives

$$\langle V_i \rangle_{i \in A} = -a \langle Q_i \rangle_{i \in A} + k_A + \langle V_i^{\mathrm{distant}} \rangle_{i \in A}, \tag{3.57}$$

where we have exploited the linearity of the mean (Eqn. (3.17)). Using the fact that $\langle Q_i \rangle_{i \in A} = Q_{A0}$ for a random alloy, which was shown earlier, this becomes

$$\langle V_i \rangle_{i \in A} = -a Q_{A0} + k_A + \langle V_i^{\mathrm{distant}} \rangle_{i \in A}. \tag{3.58}$$

Substituting Eqn. (3.40) into this gives

$$\langle V_i \rangle_{i \in A} = V_{A0} + \langle V_i^{\mathrm{distant}} \rangle_{i \in A}. \tag{3.59}$$

To show that $\langle V_i \rangle_{i \in A} = V_{A0}$, all that remains is to show that $\langle V_i^{\mathrm{distant}} \rangle_{i \in A} = 0$.

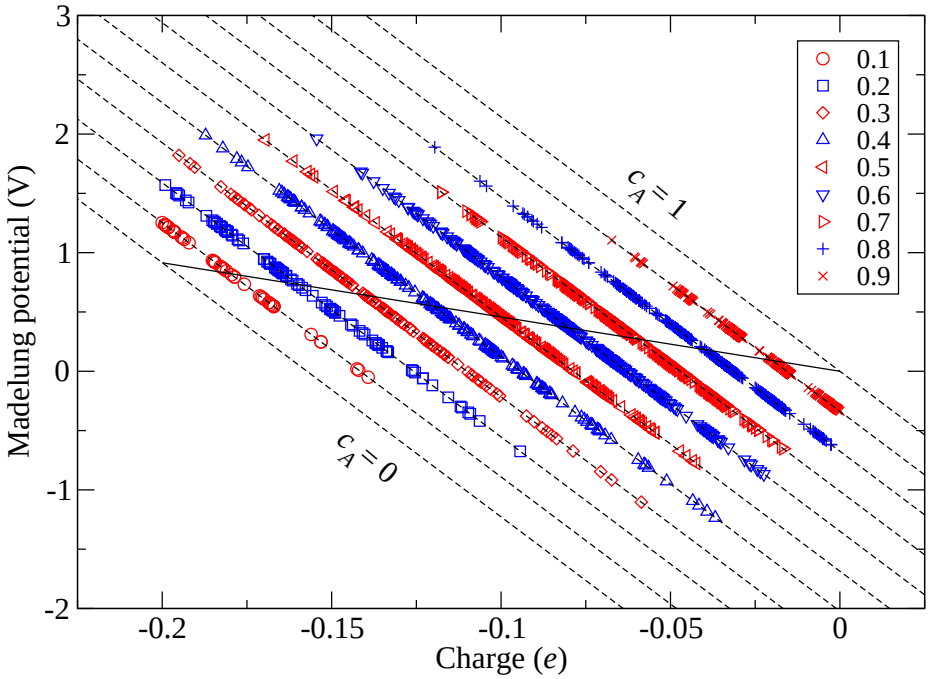


Figure 3.2: V_i vs. Q_i for fcc random alloys of various concentrations c_A. The legend indicates to which value of c_A each symbol corresponds. The dashed line which any particular set of sites lies upon is the $\langle V_i \rangle_{i \in A\,,\, Q_i = Q}$ vs. Q curve of Eqn. (3.45) for the corresponding concentration. The lowest dashed line in the figure corresponds to the $\langle V_i \rangle_{i \in A\,,\, Q_i = Q}$ vs. Q curve for $c_A = 0$, and the highest corresponds to that for $c_A = 1$. The solid line connects the points (Q_{A0}, V_{A0}) for all concentrations between $c_A = 0$ and $c_A = 1$.

$\langle V_i^{\text{distant}} \rangle_{i \in A}$ can be expressed as follows:

$$\langle V_i^{\text{distant}} \rangle_{i \in A} = \sum_Q P(Q) \langle V_i^{\text{distant}} \rangle_{i \in A\,, Q_i = Q}, \tag{3.60}$$

where the summation is over all possible charges Q which an A site can have, and $P(Q)$ is the probability of an A site chosen at random having charge Q. Using Eqn. (3.48), which applies for random alloys, the right hand side of the above equation vanishes, giving the desired result. ■

As can be seen from the figure, the (Q_i, V_i) points obtained from the calculations form a series of lines with equal gradients, and whose intercepts are linear in c_A - all of which are in excellent agreement with the analytical predictions, i.e. the $\langle V_i \rangle_{i \in A\,, Q_i = Q}$ vs. Q curves. In fact, the scatter of the (Q_i, V_i) points around their corresponding $\langle V_i \rangle_{i \in A\,, Q_i = Q}$ vs. Q curves is too small to be seen on the figure, which implies that the distant potentials are negligible for each site. The reason for this can be seen from Table 3.1: the magnitudes of λ_β are well-converged to zero by $\beta = \beta_{\max}$, and hence, for reasons described earlier, the distant potentials for each site will be negligible. In fact, for *all* bulk systems which we have considered (including the embedded thin film systems which we will discuss in a moment), the magnitudes of V_i^{distant} are all less than 0.01 V - which counts as negligible for our purposes. This means that, for practical purposes, the Q-V relations can be considered to hold here. By contrast, for smaller values of $\beta_{\max}$, it turns out that the values of λ_β are not as well converged to zero by $\beta = \beta_{\max}$. Choosing the $\boldsymbol{\lambda}$ with the highest eigenvalue and normalised such that $Q_{A0} = -0.1e$: for $\beta_{\max} = 1$ we find that $\lambda_1 = 0.008333e$; for $\beta_{\max} = 2$ we find that $\lambda_1 = 0.007667e$ and $\lambda_2 = 0.001333e$; and for $\beta_{\max} = 3$ we find that $\lambda_1 = 0.007149e$, $\lambda_2 = 0.001421e$ and $\lambda_3 = 0.0002365e$. This leads to larger values of V_i^{distant}, as can be seen in Fig. 2.4: the scatter in the values of V_i for each possible Q is larger in the fcc panel for the smaller values of $\beta_{\max}$, in which convergence of λ_β to zero by $\beta = \beta_{\max}$ is poorer.

Fig. 3.2 also illustrates that the compositions of a site's local and global environments have opposite effects on its Madelung potential. Recall that Q_i depends on the local environment of i. If the local environment consists entirely of 'like' sites, then $Q_i = 0$. Furthermore, the more 'unlike' the local environment is, then the higher the magnitude of Q_i, since site i will transfer more charge to its local environment if the local environment contains more unlike sites. Hence Q_i is a measure of how unlike the local environment of site i is.[6] With this in mind, we see from Fig. 3.2

[6] In coming to this conclusion, we have ignored the possibility that the direction of charge transfer between a pair of unlike sites depends on their separation. This is the case for the $\boldsymbol{\lambda}$ used in the

that if one increases the 'unlikeness' of a particular A site's local environment in an alloy with a particular concentration, i.e. if we make the site's charge more negative, then the result is that we increase its Madelung potential. This corresponds to moving 'up and left' along any one of the dashed lines in Fig. 3.2. On the other hand, if one increases the unlikeness of a particular A site's *global* environment while keeping its local environment fixed, i.e. if we decrease c_A while keeping Q_i constant, then the result is that we decrease the site's Madelung potential. This corresponds to moving directly downwards in Fig. 3.2. *Thus increasing the unlikeness of a site's local and global environments result in oppositely signed shifts in its Madelung potential.* We emphasise that this result is not specific to random alloys; it applies to all bulk systems which share the same $\boldsymbol{\lambda}$ and underlying lattice.

The simulated spectra for $c_A = 0.2$, 0.5 and 0.8 are shown in Fig. 3.3, where they are compared to the spectrum for a pure A metal. The pure metal spectra in the figure have all been shifted for ease of comparison. The disorder broadening can be seen clearly: the random alloy spectra are significantly wider than the pure metal spectrum. Furthermore, the disorder broadening appears to be larger for $c_A = 0.5$ than for $c_A = 0.2$ and $c_A = 0.8$. This is confirmed if one makes a quantitative comparison between the magnitudes of the disorder broadening at each concentration. This is done in Table 3.2, which gives the mean and FWHM of the distribution of $\Delta E_i^{\mathrm{B,i}}$ for A sites in each of the random alloys. The spectra can be rationalised by breaking them down into their contributions from A sites with particular local environments. Assuming that $V_i^{\mathrm{distant}} = 0$, whose validity here has been discussed earlier, and substituting Eqn. (3.42) into Eqn. (3.54), we find that

$$\Delta E_i^{\mathrm{B,i}} = \left(\frac{2}{R_1} - a\right) Q_i + k_A \tag{3.61}$$

for an A site i. Therefore, like V_i, $\Delta E_i^{\mathrm{B,i}}$ is a linear function of Q_i. In fact, they differ only in the coefficient of Q_i: for V_i the coefficient is $-a$; for $\Delta E_i^{\mathrm{B,i}}$ it is $(2/R_1 - a)$. A plot of $\Delta E_i^{\mathrm{B,i}}$ vs. Q_i analogous to Fig. 3.2 would therefore look the same, but with the data points instead lying upon lines with gradient $(2/R_1 - a)$. For the parameters used in the calculations, $(2/R_1 - a) = -11.84$ V/e. This is smaller in magnitude than a (see Table 3.1), and hence the dependence of $\Delta E_i^{\mathrm{B,i}}$ on local environment is weaker than V_i. Now, as can be seen from Table 3.1, λ_1 is significantly larger than the other values of

calculations: for A and B sites separated by R_4, the A site receives (positive) charge from the B site, while for A and B sites separated by less than R_4 the A site loses charge to the B site. This can be seen by recalling that for an A and B site separated by R_β, the A site receives charge $-2\lambda_\beta$ from the B site (where $1 \leq \beta \leq \beta_{\max}$), and by examining the elements of $\boldsymbol{\lambda}$ used in the calculations, which are given in Table 3.1. However, such anomalous charge transfer is unimportant here given the small magnitude of λ_4.

c_A	Mean (eV)	FWHM (eV)
0.2	-0.77	0.53
0.5	-0.52	0.67
0.8	-0.21	0.54

Table 3.2: Mean and FWHM of the CLS distribution for species A in various random alloys.

λ_β. Therefore the value of Q_i, and hence $\Delta E_i^{\mathrm{B,i}}$, is largely determined by the number of unlike neighbours in shell 1. We can therefore make the following approximation:

$$\Delta E_i^{\mathrm{B,i}} \approx \Delta E^{\mathrm{B,i}}(N_1), \tag{3.62}$$

where $\Delta E^{\mathrm{B,i}}(N_1)$ is the mean CLS of all A sites with $N_{i1} = N_1$. In other words, for only a small loss in accuracy we can characterise each site's local environment by the single parameter N_{i1}, instead the set of parameters $N_{i1}, N_{i2}, \ldots, N_{i\beta_{\max}}$. It turns out that the correct expression for $\Delta E^{\mathrm{B,i}}(N_1)$ is

$$\begin{aligned}\Delta E^{\mathrm{B,i}}(N_1) =& 2S_A\left(\frac{2}{R_1} - a\right)\lambda_1 N_1 \\ &+ 2S_A(1-c_A)\left[\frac{2}{R_1}\sum_{\gamma=1}^{\beta_{\max}}\lambda_\gamma Z_\gamma - \left(\frac{2}{R_1} - a\right)\lambda_1 Z_1 - \sum_{\gamma=1}^{\beta_{\max}}\frac{\lambda_\gamma Z_\gamma}{R_\gamma}\right]\end{aligned} \tag{3.63}$$

for random alloys.

Proof: We will now derive Eqn. (3.63). Consider Eqn. (3.61). Taking the mean of this over all A sites with $N_{i1} = N_1$ gives

$$\Delta E^{\mathrm{B,i}}(N_1) = \langle \Delta E_i^{\mathrm{B,i}} \rangle_{i\in A, N_{i1}=N_1} = \left(\frac{2}{R_1} - a\right)\langle Q_i \rangle_{i\in A, N_{i1}=N_1} + k_A, \tag{3.64}$$

where we have exploited the linearity of the mean (Eqn. (3.17)). We will now calculate $\langle Q_i \rangle_{i\in A, N_{i1}=N_1}$. Separating out the $\beta = 1$ term from the summation in Eqn. (2.64), and taking the same mean as above, we find that

$$\langle Q_i \rangle_{i\in A, N_{i1}=N_1} = 2S_A\lambda_1\langle N_{i1}\rangle_{i\in A, N_{i1}=N_1} + 2S_A\sum_{\beta=2}^{\infty}\lambda_\beta\langle N_{i\beta}\rangle_{i\in A, N_{i1}=N_1}, \tag{3.65}$$

where we have again exploited the linearity of the mean. Now, in a random alloy, for similar reasons as described below Eqn. (3.18),

$$\langle N_{i\beta}\rangle_{i\in A, N_{i1}=N_1} = Z_\beta c_B \tag{3.66}$$

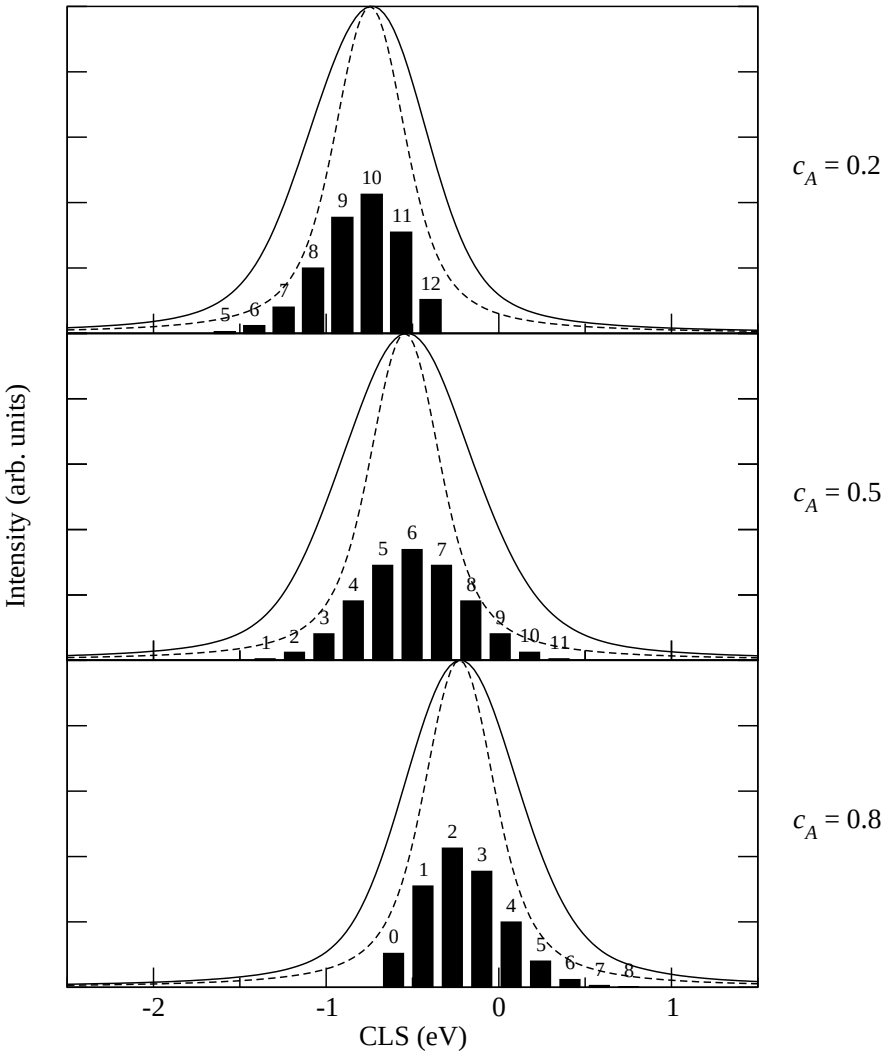


Figure 3.3: Spectra for species A in various bulk random alloys. The top, middle and bottom panels contain results pertaining to random alloys with concentrations $c_A = 0.2$, 0.5 and 0.8 respectively. In each panel, the solid curve is the spectrum for the random alloy, and the dashed curve is the spectrum for a pure A metal ($c_A = 1$) - which has been shifted for ease of comparison with that of the random alloy. The solid bars in each panel represent the contribution to the spectrum from sites with each value of N_1. The value of N_1 to which each bar corresponds is indicated. For the bar corresponding to a particular value of N_1, its height reflects the relative proportion of A sites with $N_{i1} = N_1$, and its abscissa is $\Delta E^{\mathrm{B,i}}(N_1)$.

for $\beta \geq 2$. Using this, and the fact that $\langle N_{i1} \rangle_{i \in A, N_{i1}=N_1} = N_1$, the above expression for $\langle Q_i \rangle_{i \in A, N_{i1}=N_1}$ becomes

$$\begin{aligned} \langle Q_i \rangle_{i \in A, N_{i1}=N_1} =& 2S_A \lambda_1 N_1 + 2S_A(1-c_A) \sum_{\beta=2}^{\infty} \lambda_\beta Z_\beta \\ =& 2S_A \lambda_1 N_1 + 2S_A(1-c_A) \sum_{\beta=1}^{\infty} \lambda_\beta Z_\beta - 2S_A(1-c_A)\lambda_1 Z_1, \end{aligned} \tag{3.67}$$

where we have used the fact that $c_B = 1 - c_A$. Substituting this and Eqn. (3.40) into Eqn. (3.64), and simplifying the resulting equation gives Eqn. (3.63). ■

In Fig. 3.3, the values of $\Delta E^{\mathrm{B,i}}(N_1)$ calculated using Eqn. (3.63) for each concentration are indicated by bars, whose height reflects the relative proportion $P(N_1)$ of A sites with the corresponding value of N_1.[7] Note that, while $\Delta E^{\mathrm{B,i}}(N_1)$ increases with N_1 for all concentrations, the value of $\Delta E^{\mathrm{B,i}}(N_1)$ for a particular value of N_1 decreases as the global concentration of unlike sites increases, i.e. as c_A decreases. This illustrates the fact that, like the Madelung potential, increasing how 'unlike' a site's local environment is affects its CLS in the opposite way to increasing how unlike its global environment is. Fig. 3.3 clearly illustrates the fact that, due to the linear relationship between $\Delta E^{\mathrm{B,i}}(N_1)$ and N_1 (Eqn. (3.63)), the *shape* of each spectrum must reflect that of the corresponding $P(N_1)$ vs. N_1 curve. The properties of the binomial distribution are such that the variance and skewness[8] of this curve are given by

$$\mathrm{Var}(N_1) = Z_1 c_B (1 - c_B) \tag{3.69}$$

and

$$\mathrm{Skew}(N_1) = \frac{1 - 2c_B}{\sqrt{Z_1 c_B (1 - c_B)}} \tag{3.70}$$

[7] It should be pointed out that the $P(N_1)$ illustrated in Fig. 3.3 pertain to *ideal* random alloys, i.e. they are given by the binomial distribution - Eqn. (3.19) with $\beta = 1$, and were not extracted from the periodic approximations of random alloys used in the computational calculations. However, the differences between the $P(N_1)$ extracted from the computational calculations and those predicted by the binomial distribution are inconsequential. For Fig. 3.6, which will appear in a moment, the $P(N_1)$ *were* extracted from the computational calculations.

[8] The skewness is a measure of the asymmetry of a distribution. For a random variable x, it is defined as

$$\mathrm{Skew}(x) = \frac{\langle (x - \bar{x})^3 \rangle}{\mathrm{Var}(x)^{3/2}}, \tag{3.68}$$

where $\bar{x}$ is the mean value of x and $\mathrm{Var}(x)$ is the variance of x. If $\mathrm{Skew}(x) = 0$ then the probability distribution function associated with x is symmetric about $\bar{x}$. If $\mathrm{Skew}(x) > 0$ then the tail of this function above $\bar{x}$ is longer than the tail below $\bar{x}$, i.e. the function is 'skewed' to the left; if $\mathrm{Skew}(x) < 0$ then the tail of the function below $\bar{x}$ is longer than the tail above $\bar{x}$, i.e. the function is skewed to the right.

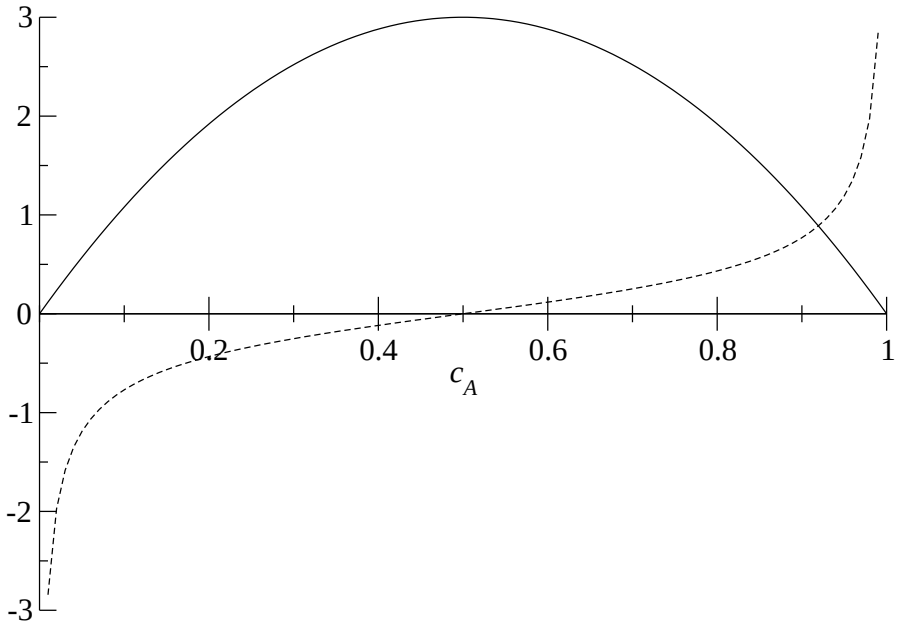


Figure 3.4: The variance (solid curve) and skewness (dashed curve) of the distribution of values of N_1 for A sites in ideal fcc random alloys as a function of c_A.

respectively [56]. These functions are plotted against $c_A = 1 - c_B$ in Fig. 3.4, where $Z_1 = 12$ was used, which corresponds to an fcc lattice, i.e. the same lattice type as was used in the calculations. It can be seen that the variance is at a maximum at $c_A = 0.5$. This is why the magnitude of the disorder broadening observed in the calculations is larger at $c_A = 0.5$ than at $c_A = 0.2$ or $c_A = 0.8$. Furthermore, the skewness is 0 at $c_A = 0.5$, and diverges at $c_A = 0$ and $c_A = 1$, with an opposite sign for $c_A > 0.5$ than for $c_A < 0.5$. This implies that the spectra will be perfectly symmetric at $c_A = 0.5$, and become increasingly asymmetric for values of c_A further away from 0.5. This effect is too small to be noticeable in the spectra for the values of c_A shown in Fig. 3.3. However, one should bear in mind the fact that asymmetries can be large if $c_A \approx 1$ or $c_A \approx 0$ when interpreting XPS spectra, in order to prevent confusion with other sources of asymmetry present in 'real' spectra, e.g. Doniach-Sunjic asymmetry. The same applies to the other systems which we will consider in this chapter, some of which exhibit very pronounced asymmetries in their spectra.

3.3.2 Embedded Thin Films

We will now consider various ordered and disordered embedded thin film (ETF) systems. Specifically, we consider individual ETFs of species A embedded in an otherwise pure B metal. We will refer to these as $B/A/B$ systems. Following Ref. [50], we modelled the concentration profile in these systems using the equation

$$c_A^l = \Lambda(l + 0.5; \sigma) - \Lambda(l - T + 0.5; \sigma), \tag{3.71}$$

where T is the total number of A layers in the system - which corresponds to the thickness of the ETF in the absence of disorder, and

$$\Lambda(l; \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{l} dt \ \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{3.72}$$

is the cumulative distribution function for the normal distribution with standard deviation σ and mean at $l = 0$. The parameter σ determines the degree of disorder in the ETF, $\sigma = 0$ being the ordered case. Larger values of σ correspond to higher degrees of diffusion of A sites from their $\sigma = 0$ positions into the surrounding B metal. To illustrate this, the values of c_A^l for $B/A/B$ systems with $T = 4$ and $\sigma = 0$, 0.75 and 1.5 - the values of σ used in the calculations - are shown in Fig. 3.5.

We considered $B/A/B$ systems with $T = 1$, 2, 4, 6, 8 and 10 and $\sigma = 0$, 0.75 and 1.5. As was mentioned earlier, the (Q_i, V_i) points in all systems were within 0.01 V of the Q-V relation for the appropriate global concentration, which is $c_A = 0$ for $B/A/B$ systems.[9] The A spectra for the systems are shown in Fig. 3.6. As was done in the case for the random alloys, for each value of N_1, the $\Delta E^{\mathrm{B,i}}(N_1)$ calculated according to Eqn. (3.63), and the relative proportion of A sites with this value of N_1, are also shown. Note that the values of $\Delta E^{\mathrm{B,i}}(N_1)$ are the same for all $B/A/B$ systems since they all share the same value of c_A (see Eqn. (3.63)). As before, the superposition of the bars for each ETF system describes the corresponding spectrum well. However, the agreement is not perfect; e.g. for the $T = 10$, $\sigma = 0$ spectrum, the peak is at a CLS lower than any possible value of $\Delta E^{\mathrm{B,i}}(N_1)$. This is due to the fact that sites in the 'bulk' of this ETF have a local environment with $N_{i\beta} = 0$ for $1 \leq \beta \leq \beta_{\max}$, while the mean local environment for a site with $N_1 = 0$ in a random alloy with $c_A = 0$ - which is what $\Delta E^{\mathrm{B,i}}(N_1)$ describes in this case - has $N_1 = 0$, and $\langle N_\beta \rangle = Z_\beta$ for $2 \leq \beta \leq \beta_{\max}$. In other words, the local environment of sites in the bulk of the ETF is more 'like' than the average local environment of A sites with $N_1 = 0$ in a random alloy with $c_A = 0$, and hence the values of $\Delta E_i^{\mathrm{B,i}}$ for the former are lower than $\Delta E^{\mathrm{B,i}}(0)$.

[9] The global concentration of A sites is 0 because the B substrate for the ETF extends infinitely in the positive and negative z directions.

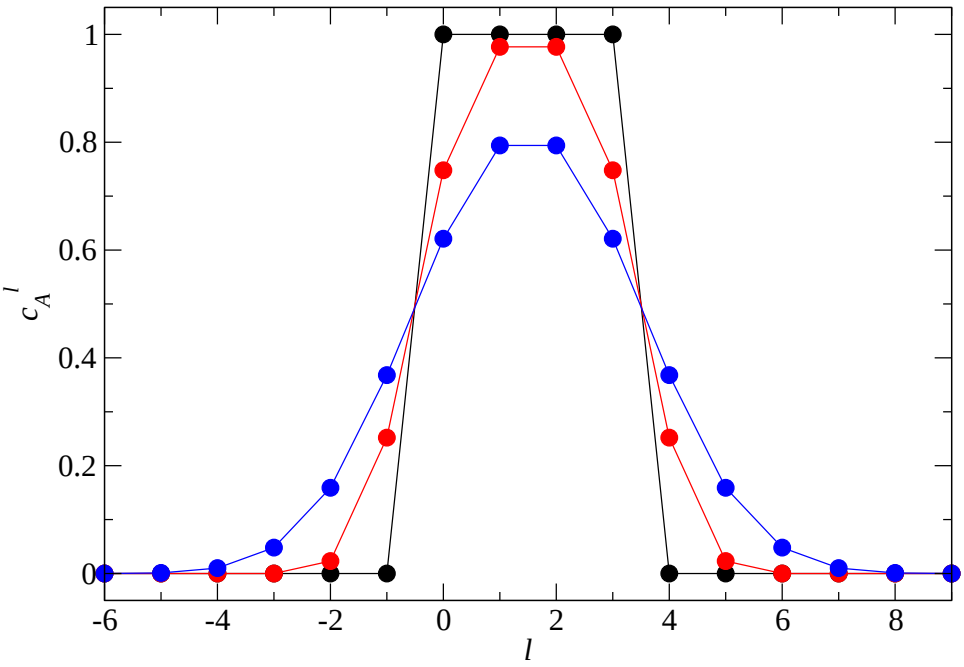


Figure 3.5: Concentration profiles for $B/A/B$ systems with $T = 4$ and $\sigma = 0$, 0.75 and 1.5, which are indicated by the black, blue and red points and lines respectively.

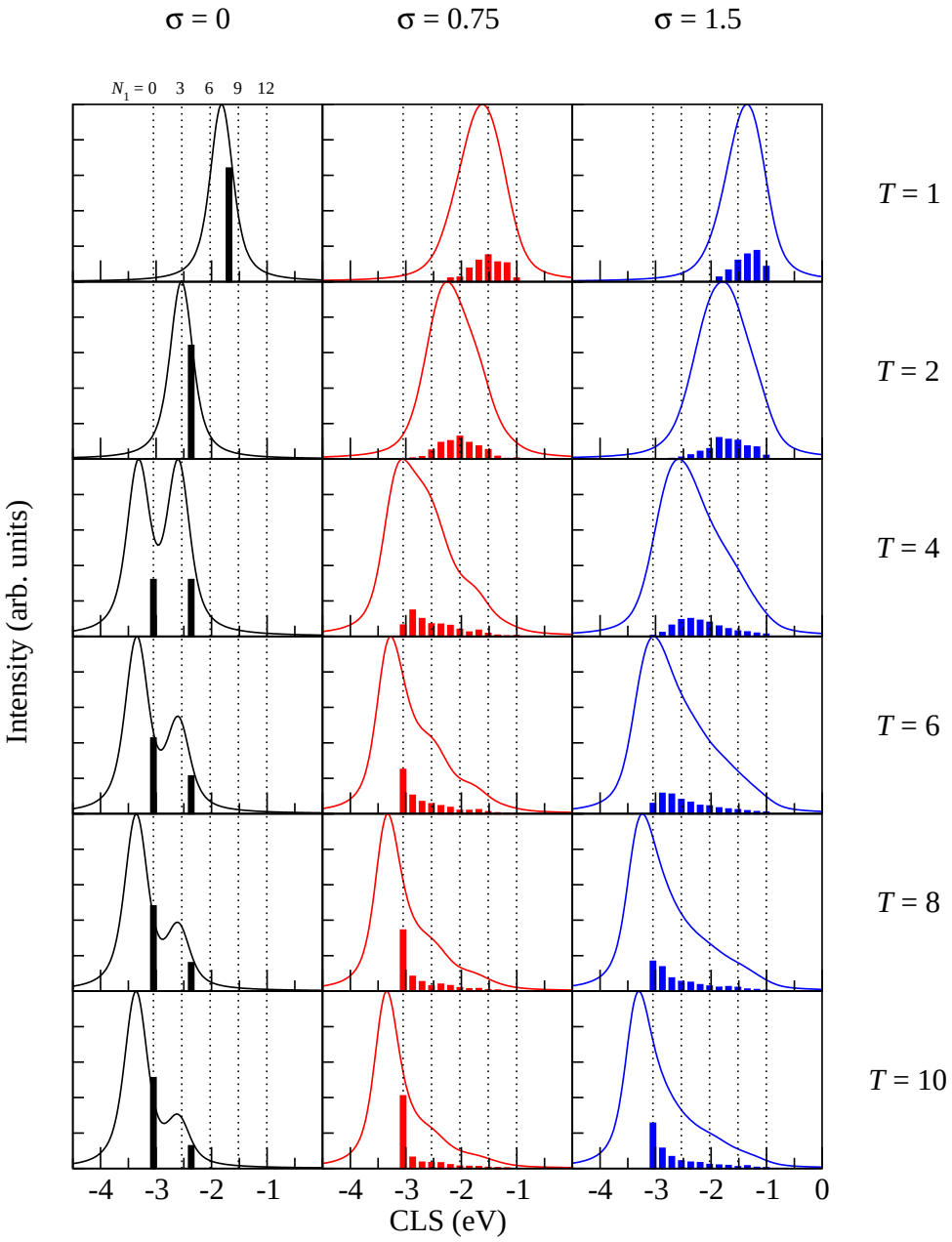


Figure 3.6: Spectra for species A in various $B/A/B$ systems. The spectra in each row correspond to ETFs with the same value of T, which is indicated to the right of the row. The spectra in each column correspond to ETFs with the same value of σ, which is indicated above each column. The dotted lines correspond to the values of $\Delta E^{\mathrm{B,i}}(N_1)$ for $N_1 = 0$, 3, 6, 9 and 12, with an arrangement indicated above the $T = 1$, $\sigma = 0$ panel. The solid bars have the same significance as described in Fig. 3.3.

However, such subtleties do not prevent us from explaining the evolution of the spectra with increasing σ in terms of the changes in the proportions of each value of N_1 within the systems. As a representative example, consider $T = 4$. At $\sigma = 0$ there are two possible local environments for A sites. Those in the middle two layers of the ETF are surrounded entirely by A sites, and therefore have $N_1 = 0$. For a site in one of the outer two layers of the ETF, 4 of its nearest neighbours are in the B region, 4 are in one of the middle two A layers, and 4 are in the same outer A layer as itself. Hence $N_1 = 4$ for such sites. This is borne out in the spectra: there are two peaks at the CLSs corresponding to $N_1 = 0$ and 4. The peaks are of equal intensity because half of the A sites are in the middle two layers, and have $N_1 = 0$; and half are in the outer two layers, and have $N_1 = 4$. Now, as σ is increased, A sites diffuse into the B region, which results in more of them having more B nearest neighbours, increasing the mean value of N_1. This manifests itself as a shift in the peak of the spectrum towards the CLS corresponding to $N_1 = 12$, as well as a widening in the spectrum on account of the larger range in local environments exhibited by the A sites. Again, this can be seen in the figure. Taking this further, we can deduce how the spectrum will continue to evolve as $\sigma \to \infty$. In the limit $\sigma \to \infty$ the A sites are completely mixed with the B metal to form a random alloy with $c_A = 0$. In this case all A sites have $N_1 = 12$, and the spectrum will therefore be narrow and centred at the CLS corresponding to $N_1 = 12$.[10] Thus, as well as the peak shifting location, the width of the spectrum will increase, and then at some point decrease, as σ is varied from 0 to ∞.

The mean and FWHM of the distribution of $\Delta E_i^{\mathrm{B,i}}$ for A sites in each of the $B/A/B$ systems are given in Table 3.3. Interestingly, the FWHM in some of these systems is more than twice as large as that observed in the $c_A = 0.5$ random alloy earlier (see Table 3.2). As mentioned earlier, the magnitude of disorder broadening is maximised in random alloys at a concentration of $c_A = 0.5$. We therefore deduce that disordered alloys with inhomogeneous concentration profiles - such as the $B/A/B$ systems considered here - can potentially exhibit disorder broadenings with magnitudes significantly larger than disordered random alloys. This effect could explain the anomalously large disorder broadening reported in Ref. [57]. In this study, it was observed experimentally that the magnitude of disorder broadening for the Ag $3d_{5/2}$ core level in various PdAg surface alloys was significantly larger than that observed in previous experimental and *ab initio* theoretical investigations of PdAg random alloys. To be more specific, PdAg alloys with Pd localised to a region near the surface, and effective concentrations of Pd ranging from $c_{\mathrm{Pd}} = 0.05$ to 0.35, exhibited disorder broadenings whose magnitude ranged from 0.4 eV to 0.66 eV. By contrast, the disorder

[10] Strictly, not all A sites will have $N_1 = 12$; however, the distribution of N_1 for A sites will have a mean of 12 and a vanishing variance.

T	σ	Mean (eV)	FWHM (eV)
1	0	-1.82	0.00
	0.75	-1.65	0.72
	1.5	-1.40	0.57
2	0	-2.54	0.00
	0.75	-2.13	0.84
	1.5	-1.77	0.89
4	0	-2.96	0.85
	0.75	-2.72	1.17
	1.5	-2.33	1.17
6	0	-3.10	0.84
	0.75	-2.93	1.19
	1.5	-2.65	1.35
8	0	-3.17	0.78
	0.75	-3.05	1.11
	1.5	-2.84	1.37
10	0	-3.21	0.72
	0.75	-3.11	1.06
	1.5	-2.94	1.31

Table 3.3: Mean and FWHM of the CLS distribution for species A in various $B/A/B$ systems.

broadenings determined theoretically and experimentally for $c_{\mathrm{Pd}} = 0.5$ random alloys are 0.35 eV [49] and 0.38 eV [46] respectively.

3.3.2.1 Comparison with *Ab Initio* Results

By reversing the signs of the CLSs of A sites in the aforementioned $B/A/B$ calculations, one has the analogous quantities for B sites in $A/B/A$ systems. We will now compare such quantities to those of Cu sites in Ni/Cu/Ni, Co/Cu/Co and Fe/Cu/Fe systems, as determined using *ab initio* calculations in Ref. [50]. The Ni/Cu/Ni and Co/Cu/Co systems considered in this study were identical to those of the model calculations, except that $R_1 = 2.49$ Å for the Ni/Cu/Ni systems and 2.47 Å for the Co/Cu/Co systems. The Fe/Cu/Fe systems differed in that, instead of a fcc crystal lattice, a bcc lattice with $R_1 = 2.45$ Å was used, though the 'layers' of the systems were still the 001 planes. The quantities which we will consider in the comparison are the *layer-resolved CLSs*. The layer-resolved CLS for B sites in layer l is defined as the mean value of $\Delta E_i^{\mathrm{B,i}}$ for such sites.

The layer-resolved CLSs of the model and *ab initio* calculations are shown in Fig.

3.7. Note that we have chosen to transform the model results such that they pertain to $A/B/A$ systems instead of $B/A/B$ systems because Cu sites in the Ni/Cu/Ni, Co/Cu/Co and Fe/Cu/Fe systems are best represented by species B in the model as opposed to species A. For comparison, the CLSs corresponding to sites with wholly like and wholly unlike local environments are indicated in the figure by dashed and dotted lines respectively. For the model results, the former quantity was calculated by setting $Q_i = 0$ - which is the charge of a site with a wholly like local environment, as discussed earlier - in Eqn. (3.61); while the latter quantity was calculated by setting $Q_i = 2S_B \sum_{\beta=1}^{\beta_{\max}} \lambda_\beta Z_\beta$ - which is the charge of a B site with a wholly unlike local environment. For the Ni/Cu/Ni *ab initio* results, the CLS corresponding to a wholly like local environment is that of a Cu site in a pure Cu metal, whose crystal structure is the same as that of pure Ni; while the CLS corresponding to a wholly unlike local environment is that of a Cu site embedded in pure Ni. The analogous quantities for the Co/Cu/Co and Fe/Cu/Fe systems were calculated similarly. It is clear from the figure that the CLS associated with each layer acts as a 'fingerprint' of the average environment of the sites within it: in the model, layers deep within the ETF, whose B sites have more 'like' environments on average, have CLSs closer to that corresponding to a wholly like environment; while those at the edge of the ETF, whose B sites have more 'unlike' environments on average, have CLSs closer to that corresponding to a wholly unlike environment. Recall that, as was described in Section 3.2, while the *ab initio* CLSs are 'true' CLSs including both initial and final state contributions - in this case determined using the total energies method described in Section 2.3.3.1 - the model CLSs consist only of the initial state contribution. Furthermore, the model CLSs for X sites are measured relative to Θ_X defined by Eqn. (3.53), for which reason it is meaningless to compare the absolute values of the *ab initio* and model CLSs. However, it is still meaningful to compare the range and shape of the *ab initio* and model CLS distributions, which we will now do.

Assuming the validity of the potential model described in Section 2.3.3.2, discrepancies between the *qualitative trends* predicted by the model and those observed in the *ab initio* results can be attributed to one or more of the following:

- The charge transfer between sites being governed by a mechanism other than that described by the OLCM.
- The values of r_i^{eff} differing between sites belonging to the same species, i.e. r_i^{eff} depending on the environment of site i.
- The dependence of $\Delta E_i^{\text{B,f}}$[11] on the environment of site i being anything other

[11] Recall that $\Delta E_i^{\text{B,f}}$ denotes the final state contribution to the CLS of site i.

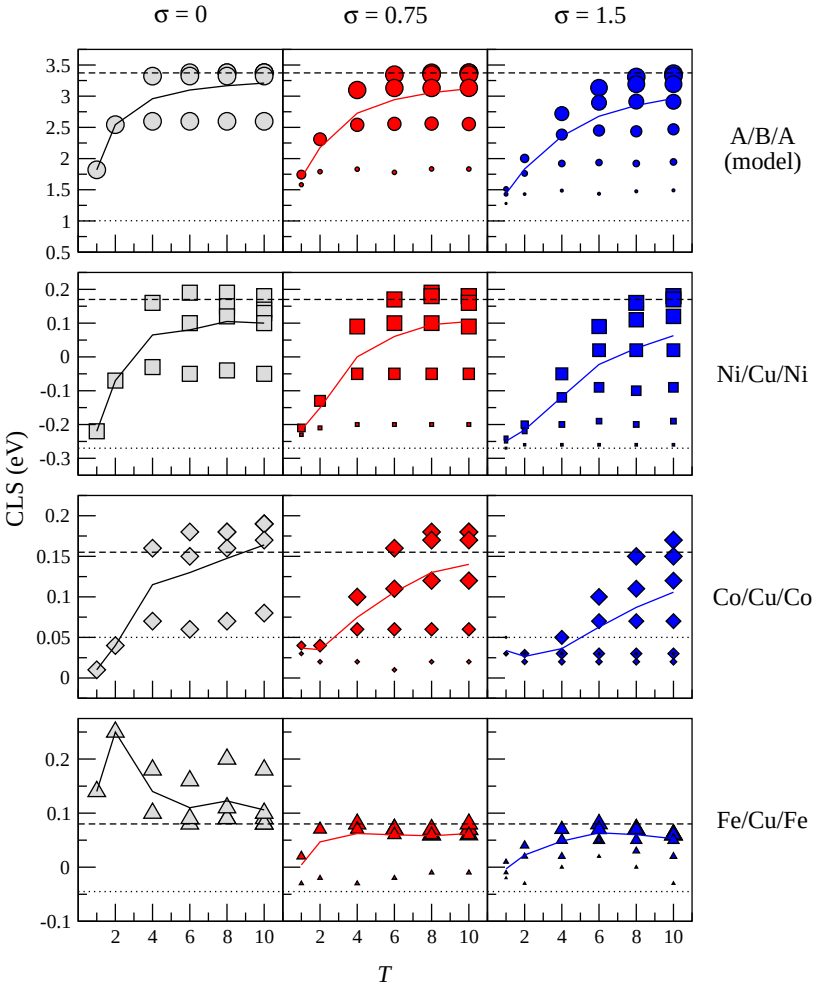


Figure 3.7: Model and *ab initio* layer-resolved CLSs for various single ETF systems. The symbols in each panel represent the layer-resolved CLS for a particular layer. The model results (top row of panels) pertain to B sites in $A/B/A$ systems. The *ab initio* results (Ref. [50]) pertain to Cu sites in Ni/Cu/Ni (second row), Co/Cu/Co (third row) and Fe/Cu/Fe (bottom row). For each row, the left, middle and right panels correspond to ETFs with $\sigma = 0$, 0.75 and 1.5 respectively. The size of each data point reflects the value of c_A^l for the layer l to which the data point corresponds to. For the model results, the dashed and dotted lines correspond to the CLS of sites with wholly like and unlike local environments respectively. For the *ab initio* results, the dashed line corresponds to the CLS of bulk Cu at Ni, Co or Fe metal volume; and the dotted line to the CLS of a Cu impurity in otherwise pure Ni, Co or Fe. The solid lines in each panel connect the mean CLS of all sites belonging to the appropriate species at the corresponding values of T.

than of the form

$$\Delta E_i^{\mathrm{B,f}} = C_X Q_i + D_X, \tag{3.73}$$

where X denotes the species of site i, and C_X and D_X are constants which, in general, depend on X and the system under consideration.

The last of the above requires clarification. In our model, $\Delta E_i^{\mathrm{B,i}}$ is given by Eqn. (3.61). Consider what the 'true' CLS of site i is if $\Delta E_i^{\mathrm{B,f}}$ has the form described by the above equation. In this case,

$$\Delta E_i^{\mathrm{B}} = \Delta E_i^{\mathrm{B,i}} + \Delta E_i^{\mathrm{B,f}} = (2/R_1 - a + C_A)Q_i + (k_A + D_A), \tag{3.74}$$

where without loss of generality we have assumed that i is an A site. Note that the above equation is of the same form as Eqn. (3.61) but with different constants; specifically, the above equation is Eqn. (3.61) with $(2/R_1 - a) \rightarrow (2/R_1 - a + C_A)$ and $k_A \rightarrow (k_A + D_A)$. Therefore, if $\Delta E_i^{\mathrm{B,f}}$ has the aforementioned form, then the *qualitative* nature of the dependence of ΔE_i^{B} with Q_i will be the same as that of $\Delta E_i^{\mathrm{B,i}}$ with Q_i. As can be seen from the figure, the model describes the qualitative trends very well for the Ni/Cu/Ni systems, and reasonably well for the Co/Cu/Co systems. From this we deduce that the effects described above collectively play a small role in the Ni/Cu/Ni systems, and a moderate role in the Co/Cu/Co systems. Further investigation is required to deduce which of the effects in particular come into play in each set of systems. In comparing the model results to those of the Fe/Cu/Fe systems it should be noted that they correspond to different crystal lattice types. However, this cannot account for such poor agreement between the two sets of results for $\sigma = 0$. Clearly the aforementioned effects become very significant here. The source of these effects was identified in Ref. [50]: the reason for the marked difference between the Fe/Cu/Fe results for $\sigma = 0$ and all other *ab initio* results is due to a well-known interface state which exists in ordered, but not in disordered, Fe/Cu/Fe systems. This accounts for the better agreement between the *ab initio* and model results for $\sigma = 0.75$ and $\sigma = 1.5$, though the agreement is still not very good. It would be interesting to see whether improved agreement is seen if a bcc crystal lattice is instead used in the model calculations.

Recall that the choice of free parameters in our calculations - which were described in Section 3.2 - do not pertain to any particular system. The next step is to repeat our calculations utilising parameters specific to the Ni/Cu/Ni, Co/Cu/Co and Fe/Cu/Fe systems. Using such system-specific parameters would presumably address the very poor *quantitative* agreement between the model and *ab initio* results: the range of model layer-resolved CLSs is more than an order of magnitude larger than that of any

of the *ab initio* results. However, in attempting to do this we run into a technical problem, which will be discussed in detail at the end of this chapter. Note that, while the quantitative agreement with the ETF *ab initio* calculations is poor, our choice of free parameters gives semi-quantitative agreement with the results of *ab initio* calculations pertaining to $Cu_{0.5}Zn_{0.5}$ random alloys. With regards to the Q-V gradients, the model gives a gradient of 21.44 V/e for both species, while the *ab initio* values are 22.68 V/e for Cu and 21.87 V/e for Zn [22, 23, 58]. With regards to the initial state contribution to the disorder broadening, the model gives a FWHM of 0.67 eV, while the *ab initio* FWHM is (for a bcc, not a fcc lattice) 0.35 eV for the Cu 2p core levels and 0.51 eV for the Zn 2p core levels [48]. This illustrates that our choice of free parameters, while inadequate with regards to the Ni/Cu/Ni, Co/Cu/Co and Fe/Cu/Fe systems, is reasonable with regards to CuZn.

Our model calculations reveal the distribution of CLSs for each species within each layer. This is not the case for the *ab initio* method utilised in Ref. [50], which only gives the mean CLS for each species within each layer. Because of this, it was necessary in that study to make an assumption regarding the width of the distribution of Cu CLSs in each layer in order to simulate the Cu XPS spectrum for each ETF system. The authors of the study assumed that the FWHM of the Cu CLSs in each layer was the same for all layers. Our model calculations can be used to test this assumption. Fig. 3.8 shows the FWHMs for B sites in each layer for the $T = 6$ $A/B/A$ systems. Also shown for comparison is the FWHM for a $c_B = 0.5$ random alloy. As can be seen from the figure, the aforementioned assumption made in Ref. [50] is unjustified: *the FWHMs can vary significantly between layers*. In fact, within one system, their values can range from 0 to the largest possible FWHM for a random alloy, i.e. a random alloy with $c_A = 0.5$.

3.4 Surface Systems

We now turn to surface systems, examining two distinct groups. Firstly, we consider random alloys. Secondly, we consider systems in which 2 layers of species A deposited upon a pure B substrate have diffused into the substrate to various degrees. For the latter group of systems, which we will refer to collectively as *surface diffusion systems*, the concentration of A sites in layer l is modelled by the formula

$$c_A^l = \Lambda(l + T + 0.5; \sigma) - \Lambda(l - T + 0.5; \sigma), \tag{3.75}$$

where $T = 4$, the parameter σ determines the degree of diffusion, and we use the following convention regarding the labelling of layers in surface systems: layer $l = 0$

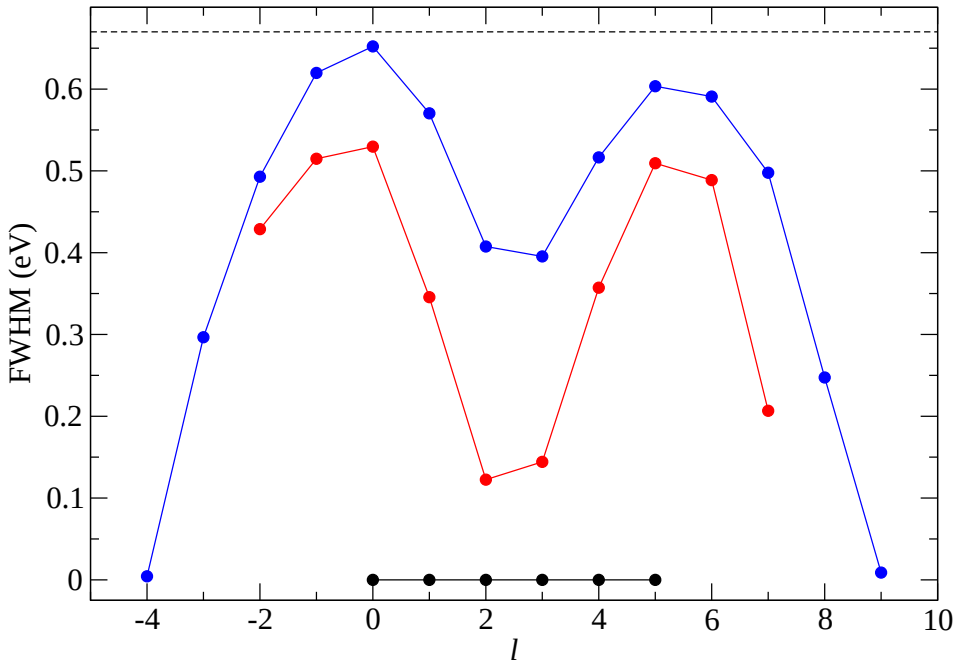


Figure 3.8: Layer-resolved FWHMs for species B in $A/B/A$ systems with $T = 6$. The black, red and blue circles connected by solid lines of the same colour correspond to the FWHMs for each layer in the $\sigma = 0$, 0.75 and 1.5 systems respectively. The dashed line corresponds to the FWHM for a random alloy with $c_B = 0.5$.

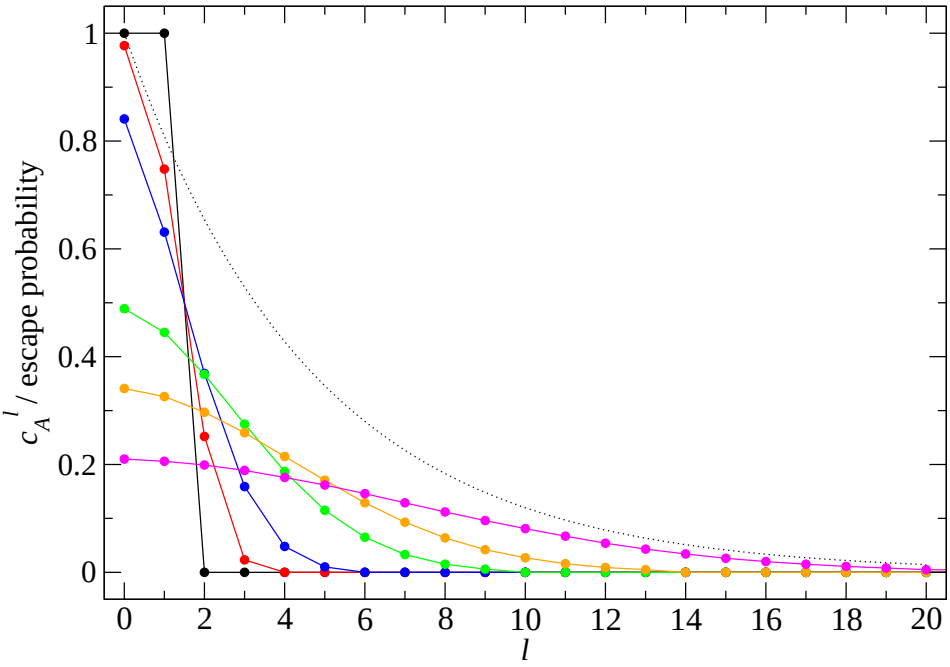


Figure 3.9: Concentration profiles for the surface diffusion systems. The black, red, blue, green, orange and magenta lines and circles correspond to the profiles for $\sigma = 0$, 0.75, 1.5, 3, 4.5 and 7.5 respectively. The dotted line connects the escape probabilities of a photoelectron with $\ell = 10$ Å created in each layer.

is the surface layer, layer 1 is the layer immediately below the surface layer, layer 2 is two layers below the surface layer, etc. By comparing the above equation to Eqn. (3.71), it can be seen that the concentration profile of a surface diffusion system with a particular value of σ is equivalent to that of 'one half' of a $T = 4$ $B/A/B$ system with the same value of σ. This can be seen by comparing Fig. 3.5 to Fig. 3.9, which shows the concentration profiles generated by the above equation for the values of σ used in the calculations: $\sigma = 0$, 0.75, 1.5, 3, 4.5 and 7.5. It should be pointed out that the validity of the OLCM in describing the values of Q_i for sites near surfaces has yet to be established. This notwithstanding, we will proceed assuming its validity. We begin by examining the relationship between V_i and Q_i in the aforementioned surface systems, and then discuss their spectra in Section 3.4.2.

3.4.1 The Relationship Between V_i and Q_i

3.4.1.1 Random alloys

Plots of V_i vs. Q_i for A sites in layers 0-5 in random alloys with $c_A = 0.2$, 0.5 and 0.8 are shown in Fig. 3.10. Also shown in the figure are the corresponding analytical predictions of Eqn. (3.45) for $\langle V_i \rangle_{i \in A\,, Q_i = Q}$ vs. Q in bulk (solid lines), and linear regression lines for the data points within each layer (dashed lines). Examination of the residuals of the linear fits (not shown) reveals that the relationship between V_i and Q_i for A sites within each layer can still be considered to be linear. However, as is clear in the figure, the linear relationship holds less rigorously than in the bulk, i.e. the scatter in the (Q_i, V_i) points for a given layer about its corresponding regression line is more pronounced. This scatter decreases as one considers layers further from the surface. From examination of the regression lines, it is clear that the gradients and intercepts associated with the linear relationship for layers 0 and 1 differ from those of layers $l \geq 2$. Layers $l \geq 2$, with regards to the V_i vs. Q_i gradients and intercepts, can be considered to be bulk-like. In layers 0 and 1, the gradient is smaller in magnitude, and the intercepts higher, than the corresponding bulk values. This is the same as was observed in Ref. [54], in which the LCM was applied to the 001 surface of a sc lattice. Interestingly, the differences between the intercepts of the regression lines for layers 0 and 1 and that of the bulk Q-V relation are smaller, the higher the concentration c_A is. Furthermore, the gradient for layer 0 is independent of c_A. The same applies to layer 1. Another noteworthy feature of Fig. 3.10 is that the magnitudes of Q_i for sites in layer 0 are generally lower than those in layers $l \geq 1$. This is expected: at the surface, sites have fewer nearest neighbours, and hence on the whole will have fewer unlike nearest neighbours, and thus a lower magnitude of charge, than sites in the bulk.

3.4.1.2 Surface Diffusion Systems

Fig. 3.11 is the analogous figure to Fig. 3.10 for the surface diffusion systems. For each system, as was the case in the random alloys, the relationship between V_i and Q_i is roughly the same for all layers $l \geq 2$, and is noticeably different for layers 0 and 1. However, as was not the case for the random alloys, the intercepts of the regression lines for layers $l \geq 2$ are not the same as the intercept of the bulk random alloy Q-V relation with the corresponding global concentration, namely, $c_A = 0$. Furthermore, the intercepts of the regression lines depend on the value of σ: the higher σ is, the closer the intercept is to that of the random alloy Q-V relation. Another difference between the random alloy and surface diffusion systems is that, in the surface diffusion systems, the gradients of the regression lines for layer 0 vary noticeably with σ. The

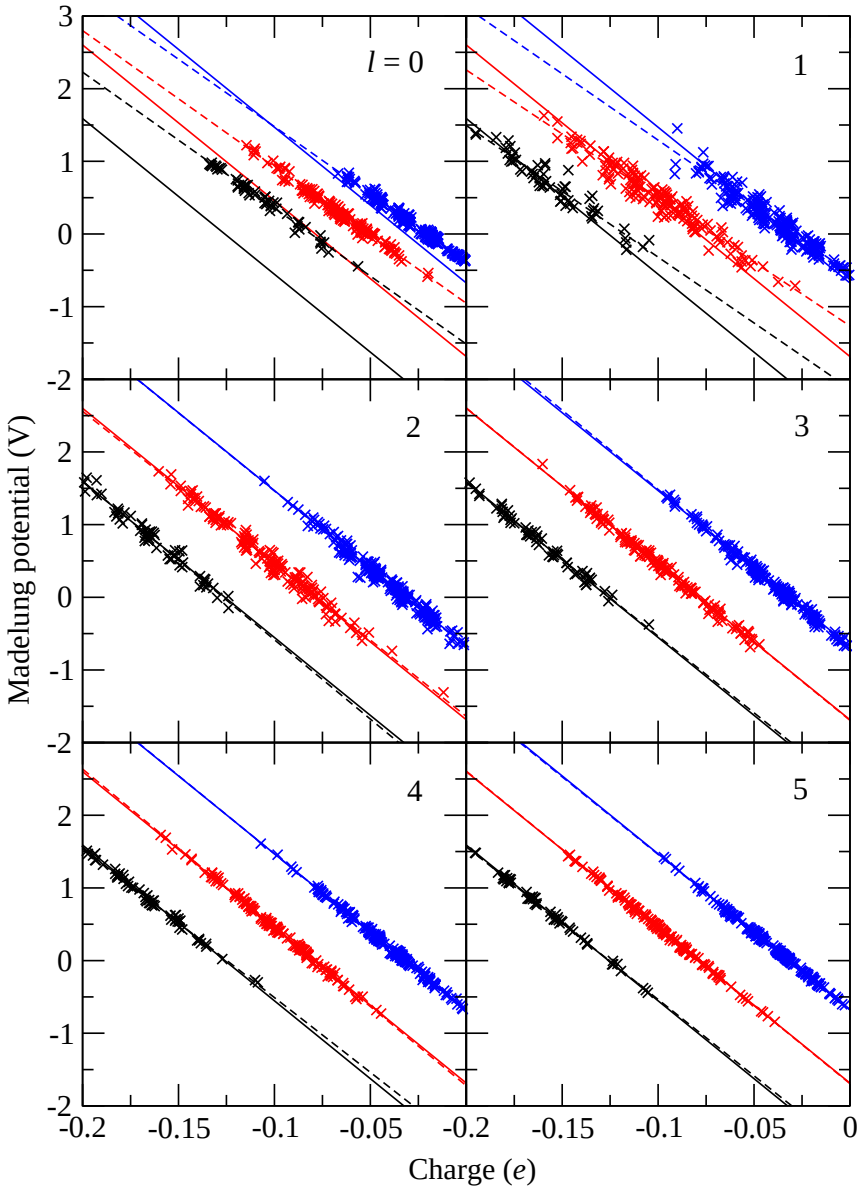


Figure 3.10: V_i vs. Q_i for A sites in layers 0-5 in random alloys with $c_A = 0.2$ (black crosses and lines), 0.5 (red) and 0.8 (blue). The data points in each panel are for a different layer, whose value of l is indicated in the upper right corner of the panel. The solid lines represent the predictions of Eqn. (3.45), while the dashed lines are a line of best fit through the corresponding data points.

same occurs for layer 1.

3.4.1.3 Analytical Results for the LCM

We will now attempt to explain some of the aforementioned trends in V_i vs. Q_i. For the sake of clarity we appeal to the LCM, i.e. the particular case of the OLCM for which $\beta_{\max} = 1$. The generalisation of what follows to any value of $\beta_{\max}$ is straightforward conceptually, though the resulting equations are considerably more complicated.

Recall that the values of Q_i in the OLCM can be constructed by adding the LCD associated with each site in turn to the system, as described in Section 3.1.1. It turns out that this procedure does *not* give the correct charges in surface systems. If this procedure were used, the LCDs for sites in layer 0 would 'spill' charge into the vacuum region. This is illustrated in Figs. 3.12(a) and 3.12(b). We can amend the procedure such that it *does* give the correct charges by modifying the LCDs for sites in layer 0 as follows. Let $\mathcal{Z}_\beta^l$ denote the number of sites in shell β of a site in layer m which are in layer n, where $|m-n| = l$. In other words, $\mathcal{Z}_\beta^l$ is the number of sites in shell β of site i which are in a single layer 'l layers away' from site i. With this in mind the *modified LCD* for an X site i, which is in layer 0, consists of a charge

$$Q'_{X0} = 2S_X(1-c_X)\lambda_1(Z_1 - \mathcal{Z}_1^1) \tag{3.76}$$

on site i itself, and a charge Q_{X1} (see Eqn. (3.2)) on each of the $Z_1 - \mathcal{Z}_1^1$ nearest neighbours of site i. The reason that site i has $Z_1 - \mathcal{Z}_1^1$ nearest neighbours instead of Z_1 is that it is 'missing' its nearest neighbours in layer -1, which resides in the vacuum region and hence contains no sites.[12] An example of a modified LCD is illustrated in Fig. 3.12(c).

Proof: We will now show that modifying the LCDs of sites in layer 0 as described above results in Eqn. (2.63) being obeyed for all sites. Consider an A site i in layer 0. The initial charge on this site is Q'_{A0}. In addition to this, the site obtains Q_{A1} from the LCDs - including modified LCDs - associated with each of its A nearest neighbours, and Q_{B1} from each of its B nearest neighbours. Therefore

$$Q_i = Q'_{A0} + N_{iA1}Q_{A1} + N_{iB1}Q_{B1}. \tag{3.77}$$

Substituting Eqns. (3.76) and (3.2) with $\beta = 1$ into the above, and simplifying the

[12] We have tacitly assumed here that nearest neighbours of a site in layer l can only reside in layers $l-1, l$ or $l+1$. This is the case if the layers are the 001 planes in an fcc lattice, as was the case in our calculations. It is also the case for the {100} planes in the sc and bcc lattices. However, it is not the case in general.

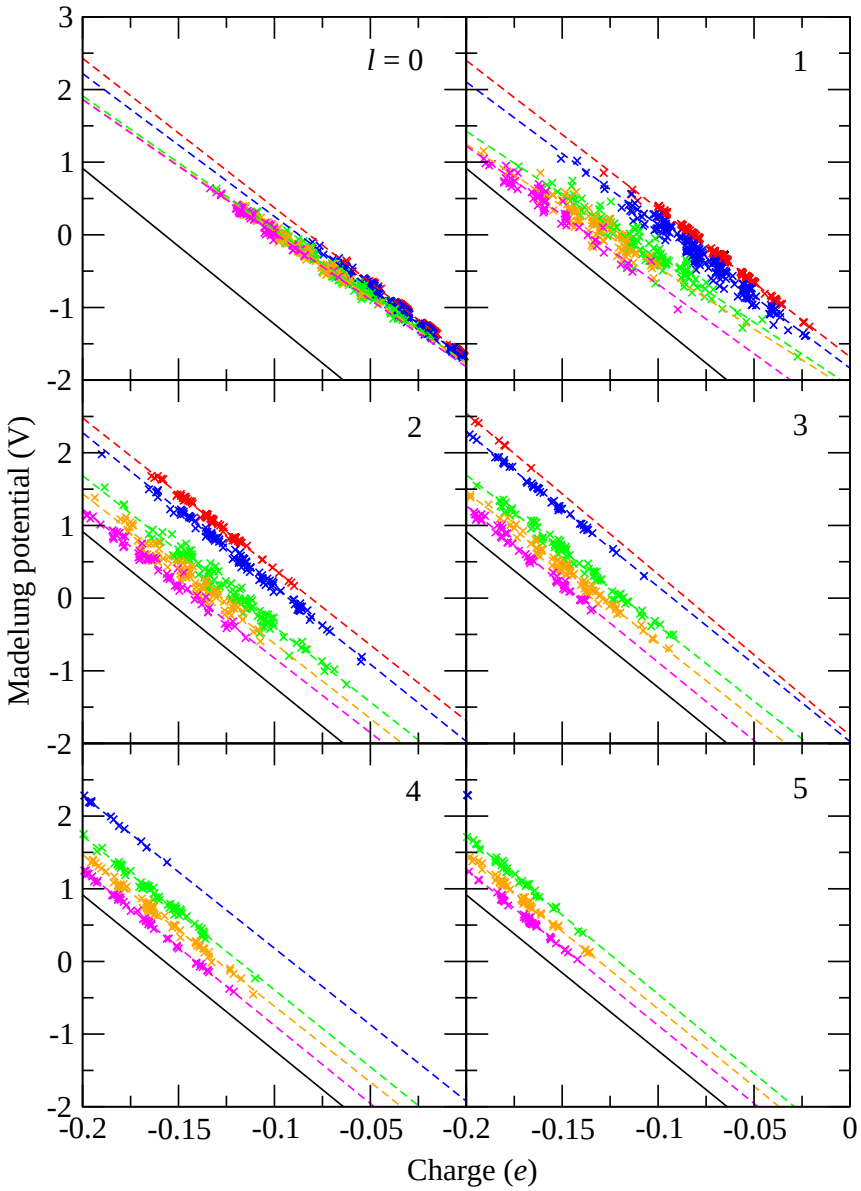
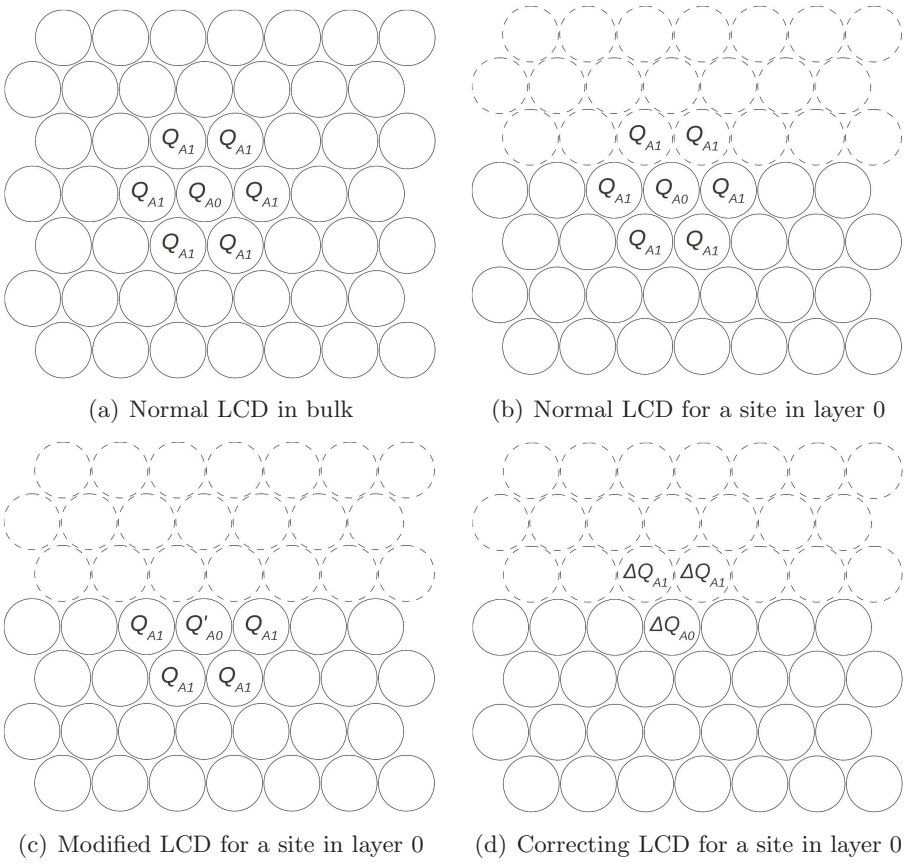


Figure 3.11: V_i vs. Q_i for layers 0-5 in the surface diffusion systems. The data points in each panel are for a different layer, whose value of l is indicated in the upper right corner of the panel. Black, red, blue, green, orange and magenta crosses and dashed lines correspond to the system with $\sigma = 0$, 0.75, 1.5, 3, 4.5 and 7.5 respectively. The dashed lines are a line of best fit through the corresponding set of crosses. The solid black line represents the predictions of Eqn. (3.45) for $c_A = 0$.



(a) Normal LCD in bulk

(b) Normal LCD for a site in layer 0

(c) Modified LCD for a site in layer 0

(d) Correcting LCD for a site in layer 0

Figure 3.12: Illustrations of the various types of localised charge distributions described in the text. The circles with solid outlines represent sites, while those with dashed outlines represent 'missing sites' which reside in the vacuum region.

resulting equation, we find that

$$Q_i = 2\lambda_1 \Big[S_A(1-c_A)(Z_1 - \mathcal{Z}_1^1) - S_A(1-c_A)N_{iA1} - S_B(1-c_B)N_{iB1} \Big]. \tag{3.78}$$

This becomes

$$Q_i = 2\lambda_1 \Big[S_A(1-c_A)(Z_1 - \mathcal{Z}_1^1) - S_A(1-c_A)(Z_1 - \mathcal{Z}_1^1 - N_{iB1}) + S_A c_A N_{iB1} \Big] \tag{3.79}$$

after noting that $S_B = -S_A$, $c_A + c_B = 1$, and $N_{iA1} + N_{iB1} = Z_1 - \mathcal{Z}_1^1$ for a site in layer 0. ($Z_1 - \mathcal{Z}_1^1$ is the total number of nearest neighbours which a site in layer 0 has). Simplifying the above equation gives

$$Q_i = 2\lambda_1 S_A N_{iB1}. \tag{3.80}$$

Noting that $N_{i1} = N_{iB1}$ for an A site, it can be seen that the above is equivalent to the LCM expression for Q_i (Eqn. (2.63) with $S_i = S_A$ since i is an A site). This derivation can be repeated with i as a B site to give an equivalent result.

For an A site i in layer $l \geq 1$, the initial charge on the site is Q_{A0}. Similarly to above, the site obtains Q_{A1} from the LCDs associated with each of its A nearest neighbours, and Q_{B1} from each of its B nearest neighbours. Therefore

$$Q_i = Q_{A0} + N_{iA1}Q_{A1} + N_{iB1}Q_{B1}. \tag{3.81}$$

Substituting Eqns. (3.1) and (3.2) with $\beta = 1$ into the above, and simplifying the resulting equation, we find that

$$Q_i = 2\lambda_1 \Big[S_A(1-c_A)Z_1 - S_A(1-c_A)N_{iA1} - S_B(1-c_B)N_{iB1} \Big]. \tag{3.82}$$

This becomes

$$Q_i = 2\lambda_1 \Big[S_A(1-c_A)Z_1 - S_A(1-c_A)(Z_1 - N_{iB1}) + S_A c_A N_{iB1} \Big] \tag{3.83}$$

after noting that $S_B = -S_A$, $c_A + c_B = 1$, and $N_{iA1} + N_{iB1} = Z_1$ for sites in layers $l \geq 1$. Simplifying the above equation, and noting that $N_{i1} = N_{iB1}$ for an A site gives the LCM expression for Q_i. ■

Note that, like the 'normal' LCDs, the modified LCDs are charge neutral. This can be seen by noting that the sum of all the charges associated with the modified LCD of an X site in layer 0 is

$$Q'_{X0} + (Z_1 - \mathcal{Z}_1^1)Q_{X1} = 2S_X(1-c_X)\lambda_1(Z_1 - \mathcal{Z}_1^1) - (Z_1 - \mathcal{Z}_1^1)2S_X\lambda_1 = 0. \tag{3.84}$$

However, unlike the normal LCDs, the symmetry of the modified LCDs is such that they have a non-zero dipole moment.

In Section 3.1 we deduced how charge on any bulk site was screened by changing the site's species and noting how the charge in the system is redistributed. We will now do the same here for a site on the surface. Let i be an A site in layer 0. Changing the species of site i to B means that the modified LCD of site i is changed from that of an A site to that of a B site. Therefore

$$\delta Q_0 = Q'_{B0} - Q'_{A0} \tag{3.85}$$

is the change in charge at site i itself, and

$$\delta Q_1 = Q_{B1} - Q_{A1} \tag{3.86}$$

is the change in charge for each of the $Z_1 - \mathcal{Z}_1$ nearest neighbours of i. Using Eqns. (3.76) and (3.2), we find that

$$\delta Q_0 = -2S_A(Z_1 - \mathcal{Z}_1^1) \tag{3.87}$$

and

$$\delta Q_1 = 2S_A\lambda_1. \tag{3.88}$$

Proof: We will now derive Eqns. (3.87) and (3.88). Substituting Eqns. (3.76) and (3.2) (with $\beta_{\max} = 1$) into Eqns. (3.85) and (3.86) gives

$$\delta Q_0 = 2\Big[S_B(1 - c_B) - S_A(1 - c_A)\Big]\lambda_1(Z_1 - \mathcal{Z}_1^1) \tag{3.89}$$

and

$$\delta Q_1 = -2\Big[S_B(1 - c_B) - S_A(1 - c_A)\Big]\lambda_1. \tag{3.90}$$

Substituting the relations $S_B = -S_A$ and $1 - c_B = c_A$ into these equations gives Eqns. (3.87) and (3.88) after simplification. ■

Using Eqns. (3.87) and (3.88), we find that the change in charge for the nearest neighbour sites, normalised to δQ_0, is

$$\phi_1 = \frac{\delta Q_1}{\delta Q_0} = -\frac{1}{Z_1 - \mathcal{Z}_1^1}. \tag{3.91}$$

By contrast, the corresponding result in the bulk is (using Eqn. (3.14) with $\beta_{\max} = 1$)

$$\phi_1 = -\frac{1}{Z_1}. \tag{3.92}$$

Since $Z_1 - \mathcal{Z}_1^1 < Z_1$, the former quantity is larger in magnitude than the latter. In other words, for a perturbation in the charge of a site i in layer 0, each nearest neighbour of i receives a larger fraction of the screening charge than would be the case if site i was in any other layer. This is due to the fact that sites in layers $l \geq 1$ are surrounded by nearest neighbours upon which screening charge can be distributed; while for sites in layer 0, there is a smaller number of nearest neighbours since it is adjacent to the vacuum region, which means that each nearest neighbour must shoulder a larger amount of the screening charge.

The modification in the LCDs of sites in layer 0 results in a shift in the Madelung potential of any site i from what would be expected by applying Eqn. (3.42). We will now derive an expression for this shift, which we will denote as ΔV_i.[13] Consider the charge distribution which must be added to the normal LCD associated with an X site j in layer 0 to transform it into the corresponding modified LCD. We will refer to this charge distribution as the *correcting LCD* of site j. As can be seen from comparing Figs. 3.12(b) and 3.12(c), this consists of a charge

$$\Delta Q_{X0} = Q'_{X0} - Q_{X0} \tag{3.93}$$

on site j itself, and a charge

$$\Delta Q_{X1} = -Q_{X1} \tag{3.94}$$

on each of its $\mathcal{Z}_1^1$ 'missing' nearest neighbour sites which are in layer -1. This is illustrated in Fig. 3.12(d). Using Eqns. (3.1), (3.76) and (3.2), the above equations become

$$\Delta Q_{X0} = -2S_X(1 - c_X)\lambda_1 \mathcal{Z}_1^1 \tag{3.95}$$

and

$$\Delta Q_{X1} = 2S_X(1 - c_X)\lambda_1. \tag{3.96}$$

If one 'normalises' the correcting LCD by dividing all of its charges by ΔQ_{X0}, the resulting charge distribution consists of a charge $+1$ on the site j itself, and a charge $-1/\mathcal{Z}_1^1$ on each of the $\mathcal{Z}_1^1$ nearest neighbours of j which are in layer -1. This follows from the above two equations. Let v_β^l be the Madelung potential at a site i in layer l which is at distance R_β from site j due to this normalised correcting LCD. Noting

[13] Note that the notation ΔV_i has a different significance in this chapter than in Chapter 2: it was used briefly in Section 2.3.3.2 to denote the extra-atomic contribution to $\Delta V_i^{\rm tot}(1)$.

that one can retrieve the un-normalised correcting LCD from the normalised correcting LCD simply by multiplying the latter by ΔQ_{X0}, it follows that the Madelung potential at site i due to the un-normalised correcting LCD is

$$\Delta V^l_{X\beta} = \Delta Q_{X0} v^l_\beta, \tag{3.97}$$

which becomes

$$\Delta V^l_{X\beta} = -2S_X(1-c_X)\lambda_1 \mathcal{Z}^1_1 v^l_\beta \tag{3.98}$$

after using Eqn. (3.95). Now, ΔV_i is the Madelung potential at site i due to the (un-normalised) correcting LCDs associated with all sites in layer 0. Therefore

$$\Delta V_i = \sum_{\beta=0}^{\infty} \left[\Delta V^l_{A\beta}\mathcal{N}_{iA\beta} + \Delta V^l_{A\beta}\mathcal{N}_{iB\beta}\right], \tag{3.99}$$

where $\mathcal{N}_{iX\beta}$ is the number of X sites in shell β of site i which are in layer 0. If we assume that i is an A site, it can be shown using Eqn. (3.98) that

$$\begin{aligned}\Delta V_i = 2S_A\lambda_1\mathcal{Z}^1_1 \Bigg\{ & v^l_1\mathcal{N}_{i1} - \left[v^l_0\mathcal{Z}^l_0 + v^l_1\mathcal{Z}^l_1\right](1-c_A) \\ & + \sum_{\beta=2}^{\infty} v^l_\beta\left[\mathcal{N}_{i\beta} - (1-c_A)\mathcal{Z}^l_\beta\right]\Bigg\},\end{aligned} \tag{3.100}$$

where $\mathcal{N}_{i\beta} = \mathcal{N}_{iB\beta}$ is the number of unlike sites in shell β of i which are in layer 0. This is the general expression for ΔV_i pertaining to an A site in layer l. Note that this depends only on the composition of layer 0. This is expected since the correcting LCDs only apply to sites in this layer.

Proof: We will now show that Eqn. (3.100) indeed follows from Eqn. (3.99). Substituting Eqn. (3.98) into (3.99) gives

$$\Delta V_i = -2\lambda_1\mathcal{Z}^1_1 \sum_{\beta=0}^{\infty} v^l_\beta\left[S_A(1-c_A)\mathcal{N}_{iA\beta} + S_B(1-c_B)\mathcal{N}_{iB\beta}\right]. \tag{3.101}$$

This becomes

$$\Delta V_i = 2S_A\lambda_1\mathcal{Z}^1_1 \sum_{\beta=0}^{\infty} v^l_\beta\left[\mathcal{N}_{iB\beta} - (1-c_A)\mathcal{Z}^l_\beta\right] \tag{3.102}$$

after using the relations $S_B = -S_A$, $c_B = 1-c_A$ and $\mathcal{N}_{iA\beta} = \mathcal{Z}^l_\beta - \mathcal{N}_{iB\beta}$, and simplifying.

Separating out the $\beta = 0$ and 1 terms, this in turn becomes

$$\Delta V_i = 2S_A\lambda_1\mathcal{Z}_1^1\Bigg\{v_0^l\left[\mathcal{N}_{iB0} - (1-c_A)\mathcal{Z}_0^l\right] + v_1^l\left[\mathcal{N}_{iB1} - (1-c_A)\mathcal{Z}_1^l\right] + \sum_{\beta=2}^{\infty} v_\beta^l\left[\mathcal{N}_{iB\beta} - (1-c_A)\mathcal{Z}_\beta^l\right]\Bigg\}. \tag{3.103}$$

Now, assuming that i is an A site, it must be the case that $\mathcal{N}_{iB0} = 0$. Substituting this into the above equation, and using the additional fact that $\mathcal{N}_{i\beta} = \mathcal{N}_{iB\beta}$ (since site i belongs to species A), Eqn. (3.100) results after simplifying. ■

We will now use Eqn. (3.100) to calculate the mean value of ΔV_i for the set of A sites in layer l which have $Q_i = Q$. We will denote this set of sites as S. Note that, from Eqn. (2.63), all sites in S must have the same value of N_{i1}, which we denote as N_1. Denoting the mean value of ΔV_i for sites in S as $\Delta V_A^l(Q)$, we find that

$$\Delta V_A^l(Q) = 2S_A\lambda_1\mathcal{Z}_1^1\Bigg\{v_1^l\langle\mathcal{N}_{i1}\rangle_{i\in S} - \left[v_0^l\mathcal{Z}_0^l + v_1^l\mathcal{Z}_1^l\right](1-c_A^0) + v^l(c_A - c_A^0)\Bigg\}, \tag{3.104}$$

where v^l is the Madelung potential of a site in layer l due to the charge distribution consisting of +1 charges being placed on all sites in layer 1 and -1 charges being placed on all sites in layer -1.

Proof: We will now derive Eqn. (3.104). Taking the mean of Eqn. (3.100) over all sites in S yields

$$\Delta V_A^l(Q) = \langle\Delta V_i\rangle_{i\in S} = 2S_A\lambda_1\mathcal{Z}_1^1\Bigg\{v_1^l\langle\mathcal{N}_{i1}\rangle_{i\in S} - \left[v_0^l\mathcal{Z}_0^l + v_1^l\mathcal{Z}_1^l\right](1-c_A) + \sum_{\beta=2}^{\infty} v_\beta^l\left[\langle\mathcal{N}_{i\beta}\rangle_{i\in S} - (1-c_A)\mathcal{Z}_\beta^l\right]\Bigg\}, \tag{3.105}$$

where we have exploited the linearity of the mean (Eqn. (3.17)). Now, for a site chosen at random from the set S, the probability of it having $\mathcal{N}_{i\beta} = \mathcal{N}_\beta$ for $\beta \geq 2$ is the same as the probability of $\mathcal{N}_\beta$ successes in $\mathcal{Z}_\beta^l$ Bernoulli trials, where each trial has a probability of success of c_B^0, i.e, the probability $P(\mathcal{N}_\beta)$ of the site having $\mathcal{N}_{i\beta} = \mathcal{N}_\beta$ is

determined by the binomial distribution:

$$P(\mathcal{N}_\beta) = \binom{\mathcal{Z}_\beta^l}{\mathcal{N}_\beta} (c_B^0)^{\mathcal{N}_\beta} (1 - c_B^0)^{\mathcal{Z}_\beta - \mathcal{N}_\beta} \tag{3.106}$$

for $\beta \geq 2$. Using the properties of the binomial distribution it therefore follows that

$$\langle \mathcal{N}_{i\beta} \rangle_{i \in S} = \mathcal{Z}_\beta^l c_B^0 = \mathcal{Z}_\beta^l (1 - c_A^0) \tag{3.107}$$

for $\beta \geq 2$. Substituting this into the above expression for $\Delta V_A^l(Q)$ gives

$$\begin{aligned} \Delta V_A^l(Q) = 2 S_A \lambda_1 \mathcal{Z}_1^1 \Bigg\{ & v_1^l \langle \mathcal{N}_{i1} \rangle_{i \in S} - \left[v_0^l \mathcal{Z}_0^l + v_1^l \mathcal{Z}_1^l \right] (1 - c_A) \\ & + \sum_{\beta=2}^{\infty} v_\beta^l \mathcal{Z}_\beta^l (c_A - c_A^0) \Bigg\} \end{aligned} \tag{3.108}$$

after simplifying. We will now *define* v^l to be

$$v^l = \sum_{\beta=0}^{\infty} v_\beta^l \mathcal{Z}_\beta^l. \tag{3.109}$$

Recall that v_β^l is the Madelung potential of a site in layer l due to a normalised correcting LCD of a site in layer 0 at distance R_β. Since there are $\mathcal{Z}_\beta^l$ possible such sites, v^l is therefore the Madelung potential of a site in layer l due to the charge distribution consisting of the sum of the normalised correcting LCDs of *all* sites in layer 0. What does this charge distribution look like? Recall the form of the normalised correcting LCD for site j: a charge of $+1$ being placed on site j itself, and a charge of $-1/\mathcal{Z}_1^1$ being placed on each of the $\mathcal{Z}_1^1$ nearest neighbours of j which are in layer -1. The charge distribution in question therefore consists of $+1$ charges being placed on all sites in layer 0, and -1 charges being placed on all sites in layer -1, where the latter follows from the fact that the normalised correcting LCDs of each of the $\mathcal{Z}_1^1$ nearest neighbours of a site k in layer -1 donates a charge $-1/\mathcal{Z}_1^1$ to site k - giving site k a total charge of -1. Hence the physical significance of v^l is as specified below Eqn. (3.104). Now, using the fact that

$$\sum_{\beta=2}^{\infty} v_\beta^l \mathcal{Z}_\beta^l = v^l - v_0^l \mathcal{Z}_0^l - v_1^l \mathcal{Z}_1^l, \tag{3.110}$$

which follows from the above definition of v^l, the above expression for $\Delta V_A^l(Q)$ becomes

$$\Delta V_A^l(Q) = 2S_A\lambda_1\mathcal{Z}_1^1\Big\{v_1^l\langle\mathcal{N}_{i1}\rangle_{i\in S} - \big[v_0^l\mathcal{Z}_0^l + v_1^l\mathcal{Z}_1^l\big](1-c_A) + \big[v^l - v_0^l\mathcal{Z}_0^l - v_1^l\mathcal{Z}_1^l\big](c_A - c_A^0)\Big\}. \tag{3.111}$$

Expanding the two sets of square brackets, and simplifying, this becomes Eqn. (3.104). ■

To proceed further, we must know the value of $\langle\mathcal{N}_{i1}\rangle_{i\in S}$. Sites in layers $l \geq 2$ have no nearest neighbours in layer 0, and hence $\mathcal{Z}_1^l = 0$ and $\mathcal{N}_{i1} = 0$ for all such sites. Noting also that $\mathcal{Z}_0^l$ is 0 for all layers l other than layer 0, we deduce that

$$\Delta V_A^l(Q) = 2S_A\lambda_1\mathcal{Z}_1^1 v^l(c_A - c_A^0) \quad \text{if } l \geq 2. \tag{3.112}$$

Note that this is independent of Q. There is therefore no correlation between ΔV_i and Q_i for sites in layers $l \geq 2$. This is why in Figs. 3.10 and 3.11 the V_i vs. Q_i regression lines for layers $l \geq 2$ are parallel to those for the analogous bulk systems. For layers deep in the bulk, i.e. at layer $l = \infty$, the above equation can be simplified further. Below we derive an expression for $\Delta V_A^\infty(Q)$ which holds if there is no 'offset' between the xy-positions of sites in adjacent layers, i.e. if for any site in any layer one can always find a site in any other layer with the same z-coordinate. For example, this is the case if the layers are the 001 planes of the sc lattice. The expression is

$$\Delta V_A^\infty(Q) = 4\pi\rho d S_A\lambda_1\mathcal{Z}_1^1(c_A - c_A^0), \tag{3.113}$$

where ρ is the number of sites per unit area in any particular layer and d is the inter-layer spacing.

Proof: We will now derive Eqn. (3.113). Recall that v^l is the is the Madelung potential of a site in layer l due to the charge distribution consisting of +1 charges being placed on all sites in layer 0 and -1 charges being placed on all sites in layer -1. This charge distribution is equivalent to a 2D lattice of dipoles occupying the $z = -d/2$ plane, where each dipole has a dipole moment $\boldsymbol{\chi}$, and $\boldsymbol{\chi}$ is the vector pointing from a -1 charge in layer -1 to one of the nearest +1 charges in layer 0. The Madelung potential of a site in layer $l = \infty$ is equivalent to the electrostatic potential at $z = \infty$ due to this charge distribution. Now, at $z = \infty$, for the purposes of calculating the electrostatic potential, the details of the charge distribution become unimportant, and we can legitimately replace the charge distribution with its continuous analogue: a plane of surface dipole

density $\boldsymbol{\tau} = \rho\boldsymbol{\chi}$ at $z = -d/2$. *If there is no xy-offset between layer -1 and layer 0*, then $\boldsymbol{\chi} = d\hat{\mathbf{z}}$, and hence $\boldsymbol{\tau} = \rho d\hat{\mathbf{z}}$. The potential at $z = \infty$ due to *this* charge distribution is equivalent to the potential at $z = \infty$ due to a plane of surface charge density $-\rho$ at $z = -d$ and a plane of surface charge density ρ at $z = 0$, which we will now evaluate.

Consider the electric field caused by the two planes. The electric field due to a single plane of charge with surface charge density ρ is of magnitude $2\pi\rho$, and is always perpendicular to the plane. With this in mind, we deduce that the electric field within the charge distribution which we are interested in is 0 for $z > 0$ and $z < -d$, and $-4\pi\rho$ in the z-direction for $-d < z < 0$. Note that the electric fields due to each plane reinforce each other in the region between the planes, and cancel outwith this region. Now, the z-component of the electric field is related to the electrostatic potential V by the equation

$$E_z = -\frac{\partial V}{\partial z}. \tag{3.114}$$

Integrating this between $z = z_1$ and $z = z_2$, it follows that

$$V(z_2) = V(z_1) - \int_{z_1}^{z_2} dz\, E_z. \tag{3.115}$$

Using the aforementioned form of the electric field, we deduce the following pair of relations:

$$V(z) = V(0) + \int_0^z dz\, 4\pi\rho = V(0) + 4\pi\rho z \quad \text{for } -d < z < 0, \tag{3.116}$$

and

$$V(\infty) = V(0). \tag{3.117}$$

Now, we can, in principle, choose the position which defines the 'zero' of the electrostatic potential to be anywhere, since only differences in electrostatic potentials have physical significance. However, our choice should be consistent with our earlier definitions of the various electrostatic potentials considered in this thesis (e.g. Hartree potential, Madelung potential), which are all of the form

$$V(\mathbf{r}) = \int d\mathbf{r}'\, \frac{q(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|}, \tag{3.118}$$

where $q(\mathbf{r})$ is some charge distribution. Applying the above equation to the current situation gives $V(\mathbf{r}) = 0$ for all positions $\mathbf{r}$ on the plane $z = -d/2$, which follows from the fact that the contribution to $V(\mathbf{r})$ at these positions due to the planes of charge at $z = 0$ and $z = -d$ exactly cancel. The zero of the electrostatic potential is therefore located at $z = -d/2$. Substituting $z = -d/2$ into the first of the above relations, and

setting $V(-d/2) = 0$, it follows that

$$V(0) = 2\pi\rho d, \tag{3.119}$$

which in turn implies that

$$V(\infty) = 2\pi\rho d \tag{3.120}$$

when substituted into Eqn. (3.117). Our earlier discussion revealed that $V(\infty) = v^{\infty}$. Therefore

$$v^{\infty} = 2\pi\rho d, \tag{3.121}$$

which, when substituted into Eqn. (3.112), gives Eqn. (3.113). ■

Eqn. (3.113) highlights the fact that that any deviation in the surface concentration c_A^0 from the global concentration c_A induces a rigid shift in the Madelung potentials of all sites deep within the bulk, which amounts to a shift in the work function of the alloy under consideration.

The situation is far more complicated for layers 0 and 1. At this point we will resort to approximations. Consider first the set of sites S pertaining to layer 0. Each of these sites has $\mathcal{Z}_1^0$ nearest neighbours in layer 0 and $\mathcal{Z}_1^1$ nearest neighbours in layer 1. The approximation which we will apply in this case is that $c_A^0 = c_A^1$. Because of this, there is no reason to believe that any one of the nearest neighbours of sites in S is more likely to be an unlike site than any other nearest neighbour. Therefore, the expected number of unlike nearest neighbours in layer 0 is

$$\langle \mathcal{N}_{i1} \rangle_{i \in S} = \frac{\mathcal{Z}_1^0}{\mathcal{Z}_1^0 + \mathcal{Z}_1^1} N_1, \tag{3.122}$$

where $\mathcal{Z}_1^0/(\mathcal{Z}_1^0 + \mathcal{Z}_1^1)$ is the fraction of nearest neighbours of a site in layer 0 which are themselves in layer 0. This would not be the case if $c_A^0 \neq c_A^1$. For example, if $c_A^0 > c_A^1$, then there is more chance of unlike nearest neighbours being found in layer 1 than layer 0, and $\langle \mathcal{N}_{i1} \rangle_{i \in S}$ will be smaller than the above value. Substituting the above expression into Eqn. (3.104), and using the fact that $\mathcal{Z}_0^0 = 1$, we find that

$$\Delta V_A^0(Q) = 2S_A\lambda_1\mathcal{Z}_1^1\left[v_1^0\frac{\mathcal{Z}_1^0}{\mathcal{Z}_1^0 + \mathcal{Z}_1^1}N_1 - (v_0^0 + v_1^0\mathcal{Z}_1^0)(1 - c_A^0) + v^0(c_A - c_A^0)\right]. \tag{3.123}$$

Expanding the brackets, and using Eqn. (2.63), this becomes

$$\Delta V_A^0(Q) = v_1^0\frac{\mathcal{Z}_1^0\mathcal{Z}_1^1}{\mathcal{Z}_1^0 + \mathcal{Z}_1^1}Q + 2S_A\lambda_1\mathcal{Z}_1^1\left[v^0(c_A - c_A^0) - (v_0^0 + v_1^0\mathcal{Z}_1^0)(1 - c_A^0)\right]. \tag{3.124}$$

Hence the mean shift for all A sites in layer 0 with charge Q is linear in Q. Note that the coefficient of Q is not the same as $-a$, and depends only on the underlying lattice. This is borne out in Fig. 3.10: the linear regression lines for layer 0 are all parallel, with a different gradient to the bulk Q-V relation. The above equation also correctly predicts the fact that the intercepts of the regression lines for layer 0 in Fig. 3.10 are closer to that of the bulk Q-V relation for higher values of c_A: $c_A^0 = c_A$ in random alloys, causing the first term in the square brackets to vanish, and resulting in a shift in the V_i vs. Q_i intercept which is proportional to $(1 - c_A)$, which tends to 0 as c_A tends to 1. The above equation therefore provides a good description of the trends observed in the OLCM calculations for random alloy surface systems. However, this is not the case for the surface diffusion systems. As can be seen from Fig. 3.11 the gradients of the regression lines vary noticeably with σ in these systems, which is at odds with the prediction of the above equation. The origin of this disagreement can be traced to the use of the assumption $c_A^0 = c_A^1$ in the derivation of the above equation, which cannot be justified for many of the surface diffusion systems.

Finally, consider the set of sites S pertaining to layer 1. Each of these sites has $\mathcal{Z}_1^1$ nearest neighbours in layer 0, $\mathcal{Z}_1^0$ nearest neighbours in layer 1, and $\mathcal{Z}_1^1$ nearest neighbours in layer 2. The approximation which we will make here is that $c_B^0 = c_B^1 = c_B^2$. In this case, similarly to the case for layer 0, we deduce that

$$\langle \mathcal{N}_{i1} \rangle_{i \in S} = \frac{\mathcal{Z}_1^1}{\mathcal{Z}_1^1 + \mathcal{Z}_1^0 + \mathcal{Z}_1^1} N_1 = \frac{\mathcal{Z}_1^1}{Z_1} N_1, \tag{3.125}$$

where we have used the fact that sites in layer 1 have a total of Z_1 nearest neighbours. Substituting the above expression into Eqn. (3.104), and using the fact that $\mathcal{Z}_0^1 = 0$, we find that

$$\Delta V_A^1(Q) = 2 S_A \lambda_1 \mathcal{Z}_1^1 \left[v_1^1 \frac{\mathcal{Z}_1^1}{Z_1} N_1 - v_1^1 \mathcal{Z}_1^1 (1 - c_A^0) + v^1 (c_A - c_A^0) \right]. \tag{3.126}$$

Again, expanding the brackets and using Eqn. (2.63), this becomes

$$\Delta V_A^1(Q) = v_1^1 \frac{(\mathcal{Z}_1^1)^2}{Z_1} Q + 2 S_A \lambda_1 \mathcal{Z}_1^1 \left[v^1 (c_A - c_A^0) - v_1^1 \mathcal{Z}_1^1 (1 - c_A^0) \right]. \tag{3.127}$$

The trends predicted by this equation are the same as those of Eqn. (3.123). Similarly to Eqn. (3.123), this equation successfully describes the trends observed in Fig. 3.10, though not those observed in Fig. 3.11.

3.4.2 Simulated Spectra

To conclude this section, we will briefly investigate what effect the aforementioned changes in the relationship between V_i and Q_i due to the presence of a surface have on the XPS spectra. Recall that our simulated spectra do not include contributions from sites in layer 0; we consider only the 'bulk' spectrum for each system.

3.4.2.1 Random Alloys

Consider random alloys first. As discussed earlier, the properties of each layer in random alloys become bulk-like by layer 2, and the properties of layer 1 are fairly close to bulk-like. We therefore expect that the presence of a surface has no discernible effect on the spectrum. The exception is if the mean free path of the photoelectrons is extraordinarily small, in which case the non-bulk-like relationship between V_i and Q_i in layer 1 may be revealed. This conclusion was also reached in Ref. [54] using the LCM.

3.4.2.2 Surface Diffusion Systems

Consider now the surface diffusion systems. The spectra for species A in these systems are shown in Fig. 3.13, where they are compared to the spectrum for species A in a random alloy with $c_A = 0$. Also shown in the figure are bars representing the contribution to each spectrum from different layers. The height of the bar corresponding to layer l is proportional to $c_A \exp(-dl/\ell)$, i.e. it is proportional to the intensity of photoelectrons originating from A sites in layer l; and the bar is centred on the mean CLS of all A sites in layer l. A quantitative description of each spectrum is provided in Table 3.4, which gives the mean and FWHM of the distribution of $\Delta E_i^{\mathrm{B,i}}$ for A sites in each of the surface diffusion systems, taking into account the fact that not all photoelectrons originating from a given site will escape the system and be detected. Specifically, the mean and FWHM values shown in Table 3.4 were calculated using the following equations:

$$\langle \Delta E_i^{\mathrm{B,i}} \rangle_{i \in A} = \frac{1}{N_A} \sum_{i \in A}{}' e^{-z_i/\ell} \Delta E_i^{\mathrm{B,i}} \tag{3.128}$$

and

$$\mathrm{FWHM} = 2\sqrt{2 \ln 2} \times \sqrt{\frac{1}{N_A} \sum_{i \in A} e^{-z_i/\ell} \Big[\Delta E_i^{\mathrm{B,i}} - \langle \Delta E_i^{\mathrm{B,i}} \rangle_{i \in A} \Big]}, \tag{3.129}$$

where N_A is the number of A sites in the system. In other words, the quantities in Table 3.4 are the mean and FWHM values which would be exhibited in the simulated spectra, if the effects of lifetime and instrumental broadening were ignored. The FWHM values are therefore a measure of the amount of disorder broadening present in the simulated

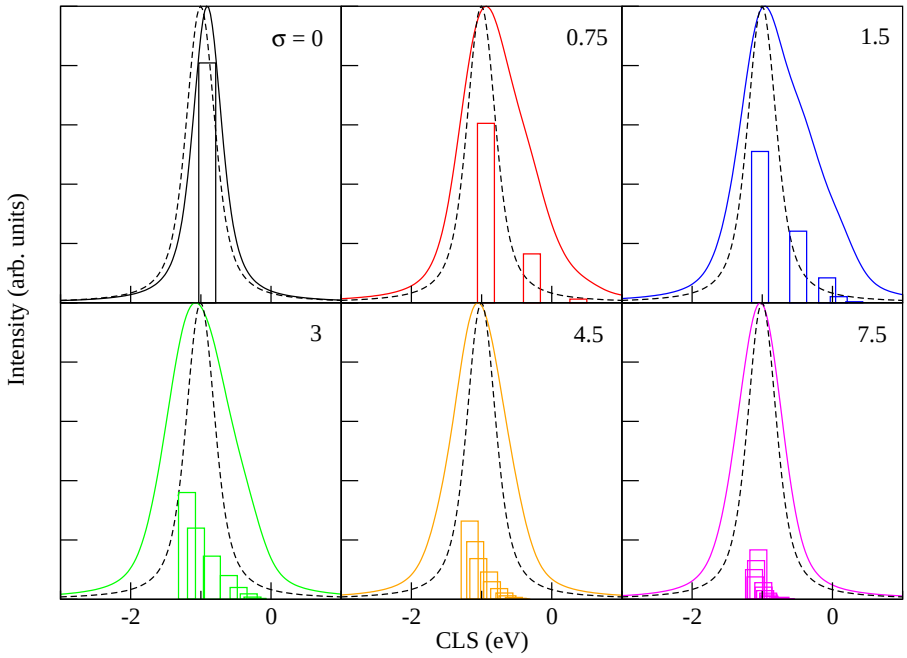


Figure 3.13: Spectra for species A in surface diffusion systems. Each panel corresponds to the surface diffusion system with a particular value of σ, which is indicated in the upper right corner of the panel. The solid curves correspond to the simulated spectra for the surface diffusion systems, while the dashed curves correspond to the bulk spectrum for species A in a random alloy with $c_A = 0$. The bars in each panel represent the contribution to the surface diffusion spectrum from A sites in a particular layer, as is described in the text.

spectra. With this in mind, note that, as was the case in the $B/A/B$ systems discussed earlier, the amount of disorder broadening is often significantly larger than is possible in a random alloy (see Table 3.2). This is further evidence of our earlier statement that inhomogeneous concentration profiles can give rise to significantly larger disorder broadenings than can be exhibited in random alloys.

We will now explain why some of the features of the spectrum vary in the manner that they do with σ. As can be seen in Fig. 3.13, the tendency is for the mean CLSs of each layer to 'bunch together' as σ is increased. This is a result of the 'flattening out' of the concentration profile as σ is increased, as can be seen in Fig. 3.9: the flatter the concentration profile, the more similar the layers become, and hence the more similar the mean CLSs of each layer become. It can also be seen in the figure, as well as in

σ	Mean (eV)	FWHM (eV)
0	-0.91	0.00
0.75	-0.78	0.91
1.5	-0.76	0.99
3	-0.99	0.84
4.5	-1.03	0.65
7.5	-1.04	0.48

Table 3.4: Mean and FWHM of the CLS distribution for species A in various surface diffusion systems.

Table 3.4, that the spectrum first increases then decreases in width as σ is increased from 0. Furthermore, the location of the peak moves very little as σ is varied. By contrast, as discussed earlier, the location of the peak in the $B/A/B$ systems moves considerably as σ was increased from 0 to ∞. Recall that the concentration profile of the surface diffusion system with a particular value of σ is the same as that of 'half' of the $T = 4$, $B/A/B$ system with the same value of σ. This means that the distribution of values of N_1 for the A sites are identical in both systems.[14] One might therefore expect the spectra in both systems to evolve in the same way as σ is increased. The reason that this does not occur can be seen in Fig. 3.11: as σ increases, there is a downward shift in the Madelung potentials of all sites in layers $l \geq 1$ in the surface diffusion systems, which seems to cancel out the expected upward shift in the peak. *Thus the presence of a surface has resulted in the spectrum evolving in a completely different manner to what would be expected from our earlier analysis of bulk systems.*

Interestingly, the evolution of the spectrum in the surface diffusion systems resembles that of the Pd $3d_{5/2}$ spectrum with time for half a layer of Pd deposited on a Ag 100 substrate at room temperature, as determined experimentally in Ref. [57]. The results of Ref. [57] are shown in Fig. 3.14. The figure shows the spectrum shortly after the Pd is deposited, 5 hours after being deposited, 15 hours after being deposited, and after annealing at 250°C. The authors of Ref. [57] interpret these results as follows. As deposited, the Pd atoms form clusters on the surface of the substrate, which are slowly annealed with time at room temperature, and the atoms eventually settle on four-fold hollow adsorption sites. Furthermore, annealing at 250°C speeds up the process. The parts of the spectrum associated with Pd atoms in the cluster and those in the adsorption sites are indicated in the figure (the notation H_4 refers to four-fold hollow adsorption sites). Note that the shoulder in the spectrum associated with the cluster slowly diminishes with time as the clusters are annealed, and disappears after annealing

[14]The presence of the surface in the surface diffusion systems is unimportant, since we are only considering sites in layers $l \geq 1$, which have no 'missing' nearest neighbours.

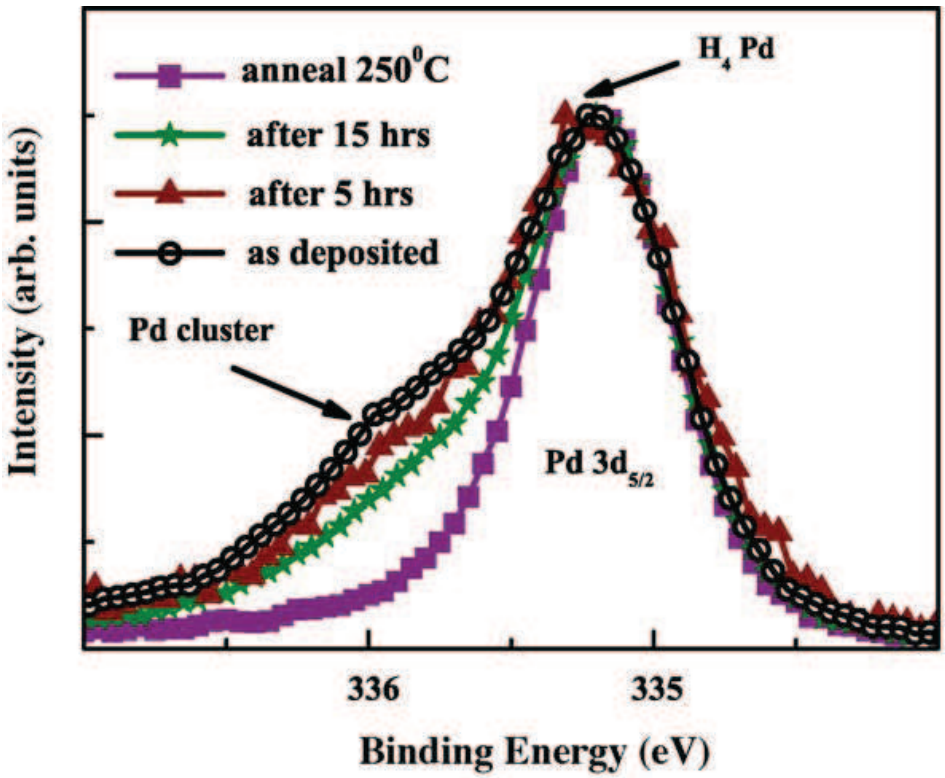


Figure 3.14: The evolution of the Pd $3d_{5/2}$ spectrum with time for half a layer of Pd deposited on a Ag 100 substrate at room temperature. (Taken from Ref. [57]).

at 250°C. Our results offer an alternative interpretation: the deposited Pd increasingly diffuses into the bulk with time. This can be seen by examining how the spectrum in Fig. 3.13 evolves as σ increases, starting at $\sigma = 1.5$, and comparing this to Fig. 3.14. It should be noted that our calculations use free parameters which do not pertain to PdAg alloys. Furthermore, our calculations are for two layers deposited upon the substrate, instead of half a layer. It would be interesting to perform OLCM calculations which can more directly compared to Fig. 3.14.

3.4.2.3 Surface Segregation

It is well known that in many alloys the concentration profile in the vicinity of a surface differs from that within the bulk [59–62]. This phenomenon is known as *surface segregation*. In addition to the random alloys and surface diffusion systems, we have also simulated spectra for various disordered alloys exhibiting surface segregation. We will refer to these systems collectively as *surface segregation systems*. The concentration profiles for these systems were modelled as being decaying and oscillatory via the

equation

$$c_A^l = (c_A^0 - c_A^\infty) \exp\left(-\frac{l}{\mu}\right) \cos\left(\frac{2\pi l}{\nu}\right) + c_A^\infty, \tag{3.130}$$

where c_A^∞ is the concentration of A sites deep within the bulk, and μ and ν are the attenuation length and wavelength of the oscillations respectively - both measured in units of the inter-layer spacing d. For all simulations we used $\mu = 2$, $\nu = 10$ and $c_A^\infty = 0.5$, and examined the following four values of c_A^0: 1, 0.8, 0.2 and 0. The corresponding concentration profiles and simulated spectra are shown in Fig. 3.15, and the mean and FWHM of the values of $\Delta E_i^{\mathrm{B,i}}$ calculated using Eqns. (3.128) and (3.129) are shown in Fig. 3.16. For comparison, the results of our earlier calculations pertaining to a random alloy bulk system with $c_A = 0.5$ are also shown in both figures. The results indicate that surface segregation induces noticeable changes in the spectrum. Firstly, there is a linear shift in the mean CLS which is proportional to $c_A^0 - c_A$. Secondly, there is an increase in the FWHM, which is larger for larger values of $|c_A^0 - c_A|$. Again, note that the inhomogeneous concentration profile results in larger magnitudes of disorder broadening than is possible in random alloys.

3.5 A Criticism of the OLCM

In this chapter we have used the OLCM to investigate the nature of the disorder broadening phenomenon in a wide range of systems. The results confirm the assertion made in Ref. [63] that disorder broadening promises to be an important tool in characterising materials on the atomic scale. The above results apply to a generic alloy consisting of species A and B. The next logical step is to apply the OLCM to a 'real' alloy. However, we run into a technical problem when we attempt to do this. Choosing the particular chemical elements to which species A and B relate amounts to choosing the vector $\boldsymbol{\lambda}$. Recall that $\boldsymbol{\lambda}$ is constrained to be an eigenvector of the f-matrix (see Eqn. (3.37)), and that its eigenvalue a turns out to be the negative of the gradient of the Q-V relations (Eqn. (3.42)). In deciding upon $\boldsymbol{\lambda}$, there are therefore two degrees of freedom: the eigenvalue of $\boldsymbol{\lambda}$, and the 'length' of $\boldsymbol{\lambda}$. Given that the f-matrix is a $\beta_{\max} \times \beta_{\max}$ matrix, it will have at most $\beta_{\max}$ distinct values of a. Ideally, one of the 'allowed' values of a would be such that it matches that of the real alloy we are trying to model. However, this is extremely unlikely. To illustrate this, the eigenvalues of the bcc f-matrix with $R_1 = 4.763a_0$[15] and the fcc f-matrix with $R_1 = 4.879a_0$ are shown for various values of $\beta_{\max}$ in Fig. 3.17. The aforementioned values of R_1 are the same as those used in the *ab initio* calculations of Ref. [23], which are for CuZn alloys. For comparison, the values of a obtained from the Q-V relations of the *ab initio*

[15]Note that a_0 denotes the Bohr radius, and has nothing to do with the eigenvalue a of the f-matrix.

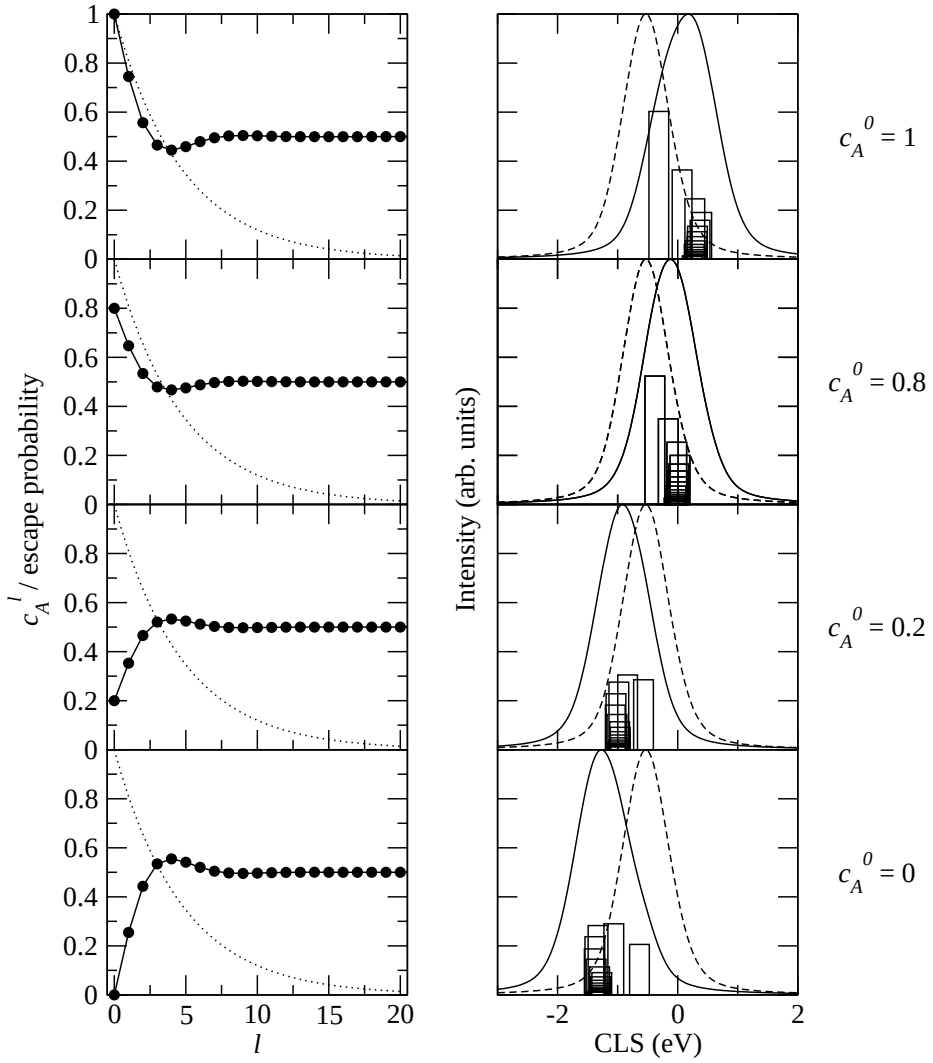


Figure 3.15: Concentration profile and spectra for species A in surface segregation systems. Each row of panels corresponds to the surface segregation system with the value of c_A^0 specified to the far right of the row. The left panel in each row gives the concentration profile, while the solid line in the right panel is the spectrum for the system. In the right column of panels, the dashed lines correspond to the bulk spectrum of a random alloy with $c_A = 0.5$, and the bars represent the contribution to the spectrum of the surface segregation system from A sites in a particular layer, as is described in the text. In the left column of panels, the dotted line connects the escape probabilities of a photoelectron with $\ell = 10$ Å created in each layer.

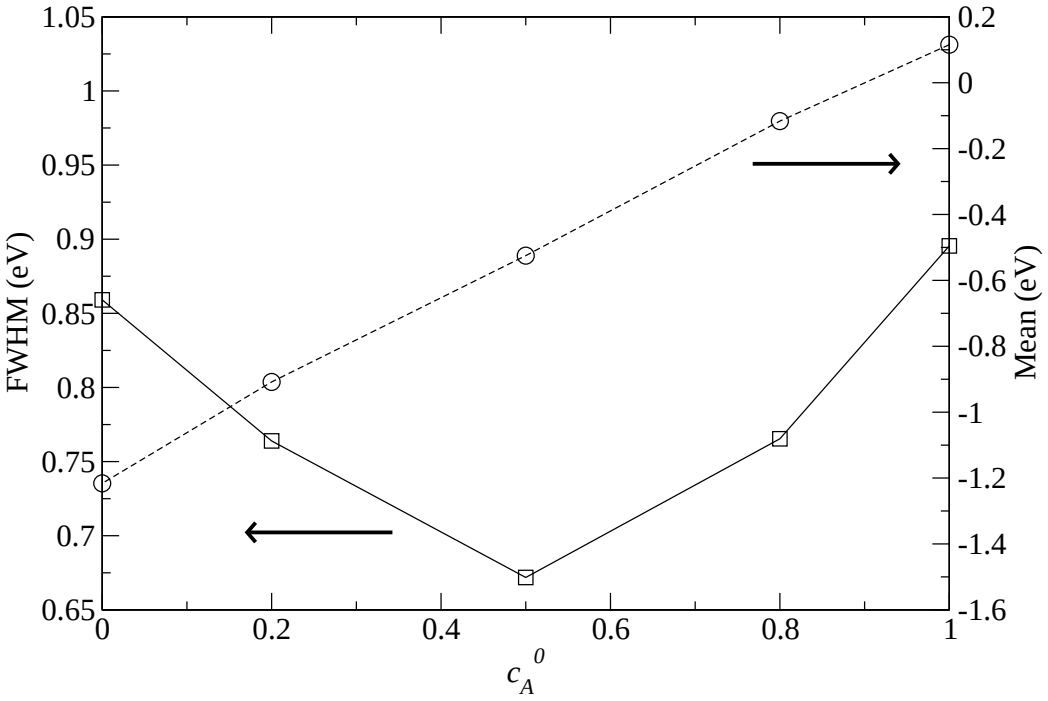


Figure 3.16: Mean and FWHM of the CLS distribution for species A in various surface segregation systems. The circles connected by the dashed lines correspond to the mean values; and the squares connected by the solid lines correspond to the FWHM values. Note that the y-axes corresponding to the mean and FWHM are different: the right y-axis corresponds to the mean values, while the left y-axis corresponds to the FWHM values - as indicated by the arrows. Note also that the results for $c_A^0 = 0.5$ pertain to a bulk calculation for a random alloy with $c_A = 0.5$.

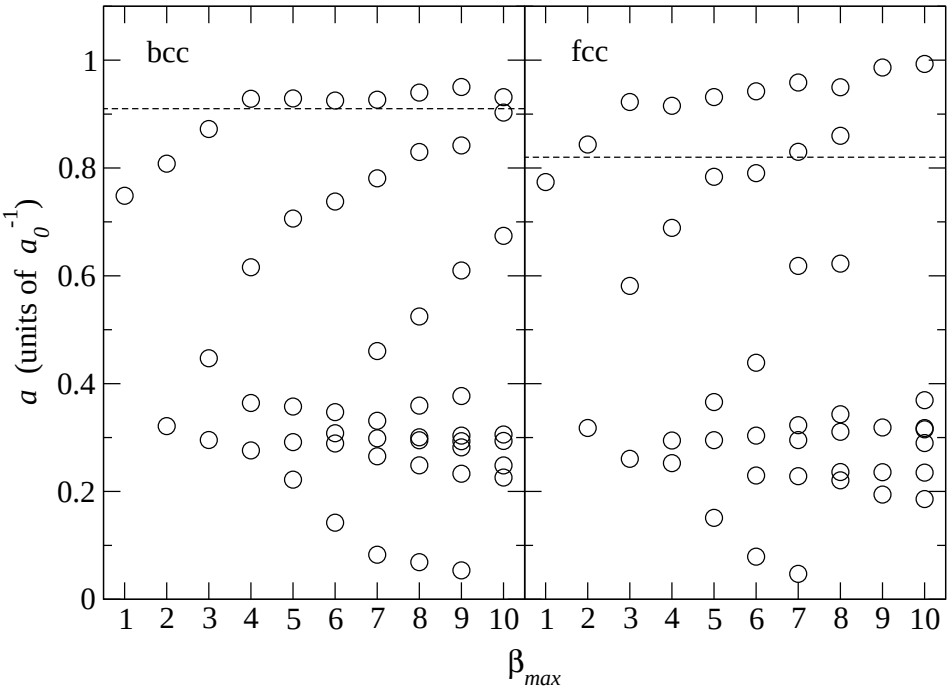


Figure 3.17: The eigenvalues a of the bcc (left panel) and fcc (right panel) f-matrices described in the text for $\beta_{\max}$ between 1 and 10. Only the real, positive eigenvalues are shown. The dotted lines in each panel correspond to the *ab initio* values of a obtained from Ref. [23].

calculations are also shown in Fig. 3.17. It is clear from the figure that, for any single value of $\beta_{\max}$, one cannot rely upon any of the allowed values of a matching the *ab initio* value, though the larger number of allowed a which comes with larger values of $\beta_{\max}$ does increase the chance of finding a value of a which is sufficiently close to the *ab initio* value to be useful. Interestingly, for the bcc lattice, one can find increasingly closer values of a to the *ab initio* value if $\beta_{\max}$ is increased from 1 to 6. This is what was observed in Ref. [26]. However, if one continues to consider values of $\beta_{\max}$ beyond 6, the closest allowed values of a to the *ab initio* value become increasingly distant from it - until one reaches $\beta_{\max} = 10$. Increasing $\beta_{\max}$ does not therefore necessarily allow better agreement with *ab initio* results. This is especially clear from the fcc panel in Fig. 3.17. The increasingly good agreement with the *ab initio* value of a observed in Ref. [26] as $\beta_{\max}$ was increased can therefore be attributed to coincidence, as well as the fact that the authors only considered $\beta_{\max} \leq 5$.

A plausible solution to the problem regarding the allowed values of a is to take the limit $\beta_{\max} \to \infty$, in which case Eqn. (2.64) becomes

$$Q_i = 2S_i \sum_{\beta=1}^{\infty} \lambda_\beta N_{i\beta}. \tag{3.131}$$

In this limit the f-matrix becomes of infinite size, and the eigenvalues form a continuum within which the exact value of a for the system under consideration can be found.[16] Setting $\beta_{\max} = \infty$ is also attractive conceptually. If $\beta_{\max} = \infty$, then the charge on any particular site depends on the composition of *all* its coordination shells - regardless of how far away they are from the site. This makes more sense than the somewhat arbitrary cut-off at shell $\beta_{\max}$ in the 'normal' OLCM. In the next chapter we examine the particular case of the CEFM (see Section 2.2.3.4) in which the values of a_i are the same for all i. This model is shown to be equivalent to the OLCM with $\beta_{\max} = \infty$ in the case of binary alloys, and can therefore be thought of as a generalisation of the OLCM which can treat multicomponent alloys - which are alloys with more than two constituent species.

3.6 Summary

To conclude this chapter, we give a summary of our key findings. The subject of this chapter was the OLCM. We began by deriving analytical results regarding the nature of the screening, and the relationship between V_i and Q_i, in the bulk of any alloy. It was pointed out that the MLCM - of which the OLCM is a particular case - is equivalent to the assumption that the screening of a charge perturbation located on any site is the same for all sites within the system under consideration.

In Section 3.3 we examined bulk systems, i.e. systems in which the effects of a surface are ignored. It was shown that the relationship between local and global environment is somewhat counterintuitive in these systems: increasing the local concentration of species A near any site i and increasing the global concentration of species A result in opposite changes to V_i, and similarly for the initial state contributions to site i's CLSs. Core level XPS spectra were then simulated for a variety of bulk systems. It was shown that the magnitude of disorder broadening in alloys with inhomogeneous concentration profiles is potentially larger - in some cases more than twice as large - than the maximum possible in analogous alloys with homogeneous concentration profiles. This was postulated as the cause of the anomalously large

[16] Strictly this is just speculation at this point: we do not yet know what the range of eigenvalues exhibited by the infinite f-matrix is. However, it is shown in the next chapter that the eigenvalues of the infinite f-matrix do indeed form a continuum.

disorder broadening reported in the experimental study of Ref. [57]. Furthermore, the simulated spectra exhibited significant asymmetries, and it was pointed out that such asymmetries - which resulted from the variety of environments in the systems - could possibly interfere with other sources of asymmetry in core level XPS spectra, e.g. Doniach-Sunjic asymmetry, causing problems with the interpretation of experimental results. The results of the simulations for ETF systems were then compared to those of analogous *ab initio* calculations pertaining to Cu ETFs embedded in Ni, Co and Fe (Ref. [50]). The model was found to describe the qualitative trends very well for the case of Ni, and reasonably well for the case of Co. The agreement was not good for the case of Fe. In all cases, the possible causes for the discrepancies between the *ab initio* and model results were identified. The assumption made in the aforementioned *ab initio* calculations that the magnitudes of disorder broadening within all layers of a given ETF system are of the same magnitude was then tested, and it was revealed that this was far from the case.

In Section 3.4 we considered surface systems, i.e. systems in which the effects of a surface are taken into account. Firstly, the status of the Q-V relations in regions near a surface was examined through the use of computational calculations, and analytically utilising the LCM - which is the simplest particular case of the OLCM. Furthermore, analytical results were derived regarding the nature of the screening in such regions. It was revealed that the Q-V relations hold to a good approximation near a surface, but with depth-dependent gradients and intercepts. It was also revealed that a difference between the global and surface layer concentrations results in a shift in the work function of the alloy under consideration. Core level XPS spectra were then simulated for various surface systems. It was shown that the surface effects are inconsequential in random alloys. However, because of the aforementioned phenomenon regarding the work function, they *are* important in systems with inhomogeneous concentration profiles near the surface. In such systems, it was shown that properly accounting for surface effects causes qualitatively different trends. These results make possible a completely different interpretation of the experimental results of Ref. [57].

Finally, a technical problem with the OLCM was highlighted in Section 3.5: the allowed values of a for a given value of $\beta_{\max}$ do not necessarily correspond to that of the specific system one is interested in. A solution to this problem was proposed.

Chapter 4

The Generalised Linear Charge Model

In this chapter we investigate a particular case of the charge-excess functional model (CEFM). For reasons which will become clear later, we will refer to this model as the *generalised linear charge model* (GLCM). We will only consider bulk systems in this chapter.

Notation

In this chapter, unless otherwise stated, we use the following notation:

- The set of sites which are in shell β of site i is denoted as β_i. Recall that shell 0 of a site i was defined in the previous chapter to consist of only site i itself.
- The number of sites in a set of sites S is denoted as $|S|$.
- The set of sites containing those sites which are in both set S *and* set T is denoted as $S \cap T$.
- The set of sites containing those sites which are in set S, excluding those sites which are also in set T, is denoted as $S \setminus T$. Similarly, the set of sites S, excluding site j, is denoted as $S \setminus j$.
- $\langle P_j \rangle_{j \in \beta_i \,, i \in T}$ denotes the mean value of P_j for all sites in shell β of each site i which is in set T
- $\langle P_j P_k \rangle_{j \in \beta_i \,, k \in \gamma_i \setminus j \,, i \in T}$ denotes the mean value of $P_j P_k$ for all pairs of sites j and k such that j is in shell β, and $k \neq j$ is in shell γ, of each site i in set T.
- The variance of the values of P_i for sites in set S is denoted $\mathrm{Var}(P_i)_{i \in S}$.

4.1 Underlying Approximations of the CEFM

Before considering the GLCM, it is useful to understand the approximations which underpin the CEFM. Since the GLCM is a particular case of the CEFM, these approximations also apply to the GLCM. To elucidate these approximations, we will now derive the CEFM energy function E defined by Eqns. (2.67), (2.66) and (2.65).

4.1.1 Derivation of CEFM Energy Function

Let

$$L_i = \int_i d\mathbf{r}\; n(\mathbf{r}) \tag{4.1}$$

denote the total number of electrons within site i, where the 'i' subscript on the integral signifies that it is over all positions within site i. Recall that we are utilising the spherical approximation (which was described at the beginning of Section 2.2), a consequence of which is that the electron density within each site is spherically symmetric about the site's nucleus. Because of this, the electron density within site i can be characterised by L_i and some function $s_i(r)$ which describes the *radial distribution* of electrons within the site. Specifically, $L_i s_i(r)$ is the electron density at distance r from $\mathbf{R}_i$, where $r \leq R_{\mathrm{WS}}$, and $s_i(r)$ is constrained to obey

$$\int_0^{R_{\mathrm{WS}}} dr\; 4\pi r^2 s_i(r) = 1 \tag{4.2}$$

such that the total electron density within site i integrates to L_i. Recall that E_L denotes the contribution to the total energy E excluding the Madelung energy, and consists of the electronic kinetic energy, the intra-site Coulomb energy and the electronic exchange-correlation energy. With the above in mind, consider the contribution to E_L from site i, which we will denote as $E_{L,i}$. To evaluate $E_{L,i}$ strictly one must know the electron density throughout the entire system, as well as the atomic numbers and positions of all nuclei. We will henceforth consider the underlying lattice to be fixed, i.e. we will not treat it as a free parameter. In this case $E_{L,i}$ is a lattice-dependent functional of the quantities z_j, L_j and $s_j(r)$ for all j. Our first assumption is that $E_{L,i}$ is a *system-dependent* functional *only* of the quantities z_i, L_i and $s_i(r)$. This can be achieved in many ways, which we will discuss later. Explicitly, this assumption is

$$E_{L,i} = \mathcal{E}_L[z_i, L_i, s_i(r)], \tag{4.3}$$

where $\mathcal{E}_L[z, L, s(r)]$ is the contribution to E_L from any site in the system under consideration which has atomic number z and L electrons with a radial distribution

described by the function $s(r)$. It is convenient at this point to define the functional

$$\mathcal{F}_i[L, s(r)] = \mathcal{E}_L[z_i, L, s(r)], \tag{4.4}$$

which we will use in a moment. Note that, since all sites belonging to the same species have the same atomic number, the functionals $\mathcal{F}_i$ are the same for all such sites.

The ground state quantities L_i and $s_i(r)$ for all i are those which minimise E subject to the following constraints: global charge neutrality, i.e.

$$\sum_i (z_i - L_i) = 0 \tag{4.5}$$

- which follows from substituting

$$Q_i = z_i - L_i \tag{4.6}$$

into Eqn. (2.68); and the validity of Eqn. (4.2) for all i. Now, the assumption made above regarding $E_{L,i}$ implies that E is given by

$$E = E_L + E_M = \sum_i E_{L,i} + E_M = \sum_i \mathcal{F}_i[L_i, s_i(r)] + E_M. \tag{4.7}$$

Note that E_M depends only on the values of L_i, and not on the functions $s_i(r)$. In other words, the inter-site Coulomb interactions 'do not care' what the radial distribution of the electron density within each site is. This is a consequence of the spherical approximation, and can be seen by substituting Eqn. (4.6) into Eqn. (2.65):

$$E_M = \frac{1}{2} \sum_i \sum_j M_{ij}(z_i - L_i)(z_j - L_j). \tag{4.8}$$

The dependence of E on $s_i(r)$ therefore enters entirely through the quantity $\mathcal{F}_i[L_i, s_i(r)]$. Because of this, the minimum in E subject to the aforementioned constraints, i.e. Eqn. (4.2) for all i and Eqn. (4.5), is equivalent to the minimum in

$$E = \sum_i F_i(L_i) + E_M \tag{4.9}$$

subject to the single constraint of Eqn. (4.5), where $F_i(L)$ denotes the minimum value of $\mathcal{F}_i[L, s(r)]$ over all $s(r)$ which obey Eqn. (4.2). The physical significance of $F_i(L)$ is as follows: $F_i(L)$ is $E_{L,i}$ if site i contains L electrons whose radial distribution is allowed to 'relax' as to obtain its minimum energy configuration. Note that, for all i, we no longer need to explicitly impose the constraint of Eqn. (4.2) because it is a

built-in feature of the function $F_i(L)$. Note also that, since the functionals $\mathcal{F}_i$ are the same for all sites belonging to the same species, then so also are the functions F_i. Now, let L_{0i} denote the value of L which minimises $F_i(L)$. We can express $F_i(L_i)$ as a Taylor series about $L = L_{0i}$ as follows:

$$F_i(L_i) = F_i(L_{0i}) + \frac{1}{2}\frac{\partial^2 F_i}{\partial L^2}(L_i - L_{0i})^2 + \frac{1}{6}\frac{\partial^3 F_i}{\partial L^3}(L_i - L_{0i})^3 + \ldots, \tag{4.10}$$

where all partial derivatives are evaluated at $L = L_{0i}$, and we have used the fact that $\partial F_i/\partial L = 0$ at $L = L_{0i}$. *Assuming that* $L_i - L_{0i}$ *is small*, then we can ignore the third and higher terms in the above, leaving

$$F_i(L_i) = F_i(L_{0i}) + \frac{1}{2}\frac{\partial^2 F_i}{\partial L^2}(L_i - L_{0i})^2. \tag{4.11}$$

Substituting this into Eqn. (4.9) gives

$$E = \sum_i F_i(L_{0i}) + \frac{1}{2}\sum_i \frac{\partial^2 F_i}{\partial L^2}(L_i - L_{0i})^2 + E_M. \tag{4.12}$$

Using Eqn. (4.6), it can easily be shown that

$$E = E_0 + \frac{1}{2}\sum_i a_i(Q_i - b_i)^2 + E_M, \tag{4.13}$$

where

$$E_0 = \sum_i F_i(L_{0i}), \tag{4.14}$$

$$a_i = \frac{\partial^2 F_i}{\partial L^2} \tag{4.15}$$

and

$$b_i = z_i - L_{0i}. \tag{4.16}$$

As can be seen from Eqns. (2.65), (2.66) and (2.67), the above expression for E is identical to that of the CEFM, except that there is additional constant E_0 added to E_L. While the extra constant is inconsequential with regards to determining the values of Q_i and V_i, it does mean that the E_L and E referred to in Section 2.2.3.4 are not 'true' local and total energies, but local and total energies *relative to* E_0. We will continue to use this convention when considering the GLCM later, i.e. from Section 4.2 onwards all local and total energies are measured relative to E_0. This point is especially important with regards to the results of Section 4.8.2, where the *composition*-dependence of E_L and E in random alloys are discussed, where by the term 'composition' we mean the

values of c_X for all X; it is important because E_0 is itself a composition-dependent quantity.

In Section 2.2.3.4 we defined the *local interactions* as those associated with E_L, and equated a_i to the strength of the local interactions which act to keep the charge of site i at its 'bare charge' b_i. We also stated that the values of a_i and b_i are the same for all sites belonging to the same species. This is all consistent with the above, as we will now show. Recall that $F_i(L)$ is the contribution to E_L from site i if the site has L electrons, and L_{0i} is the value of L which minimises $F_i(L)$. From Eqns. (4.15) and (4.16) it can be seen that b_i is the charge on site i which occurs at the minimum in $F_i(L)$, and a_i is the curvature of $F_i(L)$ at this minimum. Therefore b_i is the charge which the local interactions within site i 'want' site i to have, and a_i is the strength of these interactions. In fact, since b_i is the charge which site i 'wants' to be, one can regard b_i as a measure of the *electropositivity*, i.e. the propensity to attract (positive) charge, of site i within the system under consideration - a point which will be become important later. Furthermore, since - as was mentioned earlier - $F_i(L)$ is the same for all sites belonging to the same species, then so also are the values of a_i and b_i: $a_i = a_X$ and $b_i = b_X$ if site i belongs to species X.

4.1.2 Underlying Approximations

In deriving the CEFM energy function above the only approximations which we have made are as follows:

1. The spherical approximation.
2. That $E_{L,i}$ is a system-dependent functional only of the contents of site i.
3. That $Q_i - b_i$ is small for all i.

The last of these is equivalent to the approximation made earlier that $L_i - L_{i0}$ is small for all i - a fact which can easily be derived from Eqns. (4.6) and (4.16). As was mentioned in Section 2.2.3.4, the Q-V relations are implicit in the CEFM; they result automatically from minimising Eqn. (4.13). Therefore any model which utilises the above approximations is equivalent to the CEFM, and will exhibit the Q-V relations.

Earlier it was mentioned that the second of the above approximations - which we will henceforth refer to as the *local approximation* - can be achieved in many ways. We will now elaborate on this point. Firstly, it is the case if, for the purposes of evaluating $E_{L,i}$, the region outwith site i is approximated as an effective medium whose properties somehow reflect the system as a whole. In other words, with regards to calculating $E_{L,i}$ for each site, all sites 'see' the effective medium as their surroundings. We will refer to

this manner of achieving the local approximation as the *effective medium approach*. It is from this perspective that the CEFM energy function was derived in Ref. [33] within the framework of multiple scattering theory. Note that, if the effective medium is the same for all systems with the same underlying lattice and composition, then so also are E_0, and a_X and b_X for any particular species X, i.e. the quantities E_0, a_X and b_X are *transferable* between systems with the same underlying lattice and composition. This is the case for the SSCPA effective medium described in Section 2.2.1.

An alternative manner in which the local approximation can be achieved is if, instead of an effective medium, one assumes that the region outwith each site is a vacuum when calculating $E_{L,i}$. We will refer to this as the *vacuum approach*. In this case $E_{L,i}$ becomes equivalent to the energy required to 'charge' site i in isolation to have L_i electrons. Furthermore, the values of a_X and b_X for any particular species X will be the same for all alloys with the same underlying lattice regardless of their composition, i.e. the quantities a_X and b_X are transferable between systems with the same underlying lattice in the vacuum approach. It is by using the vacuum approach that the CEFM energy function was derived in Ref. [26].

Finally, the local approximation results from the combined use of the Thomas-Fermi and local density approximations (see Sections 2.1.2 and 2.2.3.1). For reasons which will become clear in a moment, we will refer to this as the *Pinski approach*. Let $E_{\mathrm{kin},i}$, $E_{\mathrm{intra},i}$ and $E_{\mathrm{xc},i}$ denote the contributions from site i to the kinetic, intra-site Coulomb and exchange-correlation energies of the system respectively. Note that

$$E_{L,i} = E_{\mathrm{kin},i} + E_{\mathrm{intra},i} + E_{\mathrm{xc},i}. \tag{4.17}$$

Now, within the Thomas-Fermi and local density approximations

$$E_{\mathrm{kin},i} = \int_i d\mathbf{r}\ \epsilon_{\mathrm{kin,HEG}}(n(\mathbf{r})) \tag{4.18}$$

and

$$E_{\mathrm{xc},i} = \int_i d\mathbf{r}\ \epsilon_{\mathrm{x,HEG}}(n(\mathbf{r})) + \int_i d\mathbf{r}\ \epsilon_{\mathrm{c,HEG}}(n(\mathbf{r})), \tag{4.19}$$

where recall that $\epsilon_{\mathrm{kin,HEG}}(n)$, $\epsilon_{\mathrm{x,HEG}}(n)$ and $\epsilon_{\mathrm{x,HEG}}(n)$ are the kinetic, exchange and correlation energies per unit volume respectively of a homogeneous electron gas with electron density n. Furthermore, the *exact* expression for $E_{\mathrm{intra},i}$ is

$$E_{\mathrm{intra},i} = \frac{1}{2}\int_i d\mathbf{r} \int_i d\mathbf{r}'\ \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \int_i d\mathbf{r}\frac{z_i n(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_i|}. \tag{4.20}$$

It is clear that the above three quantities depend only on the contents of site i. This largely explains why Pinski's model - which was described in Section 2.2.3.1 - reproduces

the Q-V relations. In Pinski's model all but the last of the approximations listed earlier are utilised,[1] with the local approximation being achieved through the combined use of the Thomas-Fermi and local density approximations. It therefore *must* be the case that the Q-V relations occur in Pinski's model if the values of $Q_i - b_i$ are sufficiently small. Note that, as was the case in the vacuum approach described above, in the Pinski approach the values of a_X and b_X are transferable between systems with the same underlying lattice.

4.2 Fundamental Properties of the GLCM

We now turn to the GLCM. The GLCM is the particular case of the CEFM in which $a_i = a$ for all i. The key features of the CEFM were discussed earlier in Section 2.2.3.4. Setting $a_i = a$ for all i in Section 2.2.3.4 therefore recovers the analogous features for the GLCM. These are as follows. In the GLCM, the values of Q_i are those which minimise the energy function

$$E = E_L + E_M, \tag{4.21}$$

where

$$E_L = \frac{1}{2} a \sum_i (Q_i - b_i)^2 \tag{4.22}$$

is the local energy,

$$E_M = \frac{1}{2} \sum_i \sum_j M_{ij} Q_i Q_j \tag{4.23}$$

is the Madelung energy, and the parameter b_i depends only on the species of site i, and takes the value b_X if i belongs to species X. The free parameters in the model are: b_X for each species X present in the system; and a. The physical significance of the CEFM free parameters have already been discussed in Section 2.2.3.4 and the previous section. The same applies to the GLCM, with the addendum that, since setting $a_i = a$ for all i in the CEFM recovers the GLCM, the strength of the local interactions in the GLCM is the same for all sites. In the GLCM, minimising E leads to the following equation:

$$V_i = -aQ_i + k_i, \tag{4.24}$$

where

$$k_i = ab_i. \tag{4.25}$$

[1] In Ref. [25], Pinski briefly utilised a version of his model in which the spherical approximation does not hold. However, we do not consider this here.

In other words, the Q-V relations are implicit in the GLCM: V_i and Q_i for X sites will *always* form a Q-V relation with gradient a and intercept

$$k_X = ab_X. \tag{4.26}$$

4.2.1 Constraints on the Free Parameters

It was pointed out in Ref. [32] that, for the CEFM, E has a minimum if and only if the matrix

$$H_{ij} = a_i\delta_{ij} + M_{ij} \tag{4.27}$$

is positive-definite,[2] where δ_{ij} is the Kronecker delta. Setting $a_i = a$ gives the analogous condition for the GLCM: E has a minimum if and only if the matrix

$$H_{ij} = a\delta_{ij} + M_{ij} \tag{4.28}$$

is positive-definite. Since M depends only upon the underlying lattice (see Eqn. (2.55)), and we require that E has a minimum, it follows that a is constrained such that H is positive-definite. This constraint is somewhat abstract, and does not give a feel for what values of a are 'allowed'. To rectify this matter, it can be recast as follows: a must obey the inequality

$$a > a_{\min}, \tag{4.29}$$

where $a_{\min}$ is a non-negative, finite quantity which depends only on the underlying lattice.

Proof: Consider first the Madelung matrix M, whose elements are defined in Eqn. (2.55). From Sylvester's criterion,[3] a necessary condition for M to be positive-definite is that

$$M_{11} > 0. \tag{4.30}$$

As can be seen from Eqn. (2.55), $M_{11} = 0$, and hence this condition is not met. Therefore M is not positive-definite. Since M is positive-definite if and only if all of its eigenvalues are positive, it must therefore be the case that M has at least one eigenvalue which is non-positive. Denoting the eigenvalue of M which has the lowest value as $\lambda_{\min}$, it must therefore be the case that $\lambda_{\min}$ is non-positive. We will use this result in a moment.

[2] A matrix is positive-definite if and only if all of its eigenvalues are all positive.

[3] Sylvester's criterion states that any matrix A is positive-definite if and only if the determinants of all of the upper-left submatrices (including the 1×1 submatrix A_{11}) of A are positive.

Let $\mathbf{v}$ denote an eigenvector of M with eigenvalue λ, i.e.

$$M\mathbf{v} = \lambda\mathbf{v}. \tag{4.31}$$

Premultiplying $\mathbf{v}$ by the matrix H gives

$$H\mathbf{v} = M\mathbf{v} + a\mathbf{v}, \tag{4.32}$$

where we have used Eqn. (4.28). This equation in turn becomes

$$H\mathbf{v} = (\lambda + a)\mathbf{v} \tag{4.33}$$

after using Eqn. (4.31), which reveals that: for every eigenvalue λ of M there is a corresponding eigenvalue $\lambda + a$ of H. Now, H is positive-definite if and only if all of its eigenvalues are positive. This amounts to the following condition: for every eigenvalue λ of M, a must be such that

$$\lambda + a > 0, \tag{4.34}$$

or equivalently,

$$a > -\lambda. \tag{4.35}$$

This is the case if and only if a obeys Eqn. (4.29) with

$$a_{\min} = -\lambda_{\min}. \tag{4.36}$$

Since M depends only on the underlying lattice, then so also does $\lambda_{\min}$. It therefore follows from Eqn. (4.36) that $a_{\min}$ depends only on the underlying lattice. Furthermore, since - as was deduced earlier - $\lambda_{\min}$ is non-positive, it also follows from Eqn (4.36) that $a_{\min}$ must be non-negative. All that remains is to show that $a_{\min}$ is finite. To do this, we appeal to the following property of the Madelung matrix M, which was described in Ref. [32]: for each wavevector $\mathbf{k}$ in the Brillouin zone associated with the underlying lattice, there is a corresponding eigenvalue

$$M(\mathbf{k}) = \frac{B}{|\mathbf{k}|^2} + m(\mathbf{k}), \tag{4.37}$$

of the Madelung matrix M, where B is a positive constant, $m(\mathbf{k})$ is a bounded function of $\mathbf{k}$, and B and $m(\mathbf{k})$ both depend only on the underlying lattice. Note that the first term in $M(\mathbf{k})$ diverges to $+\infty$ as $\mathbf{k} \to 0$. Since $m(\mathbf{k})$ is bounded, there is therefore no divergence in $M(\mathbf{k})$ to $-\infty$, and hence the lowest eigenvalue $\lambda_{\min}$ of M must be finite. From Eqn. (4.36) it must therefore be the case that $a_{\min}$ is finite. ■

We now turn to the values of b_i. In Section 2.2.3.4 we minimised E for the CEFM, but did not explicitly enforce the constraint of global charge neutrality. The same was done in Section 4.2 for the GLCM. This was justified in Section 2.2.3.4: it is unnecessary to explicitly constrain the system to be charge neutral when minimising E since the system is automatically charge neutral at its minimum [32]. However, if one does not include the charge neutrality constraint when minimising E, then, for the GLCM, it is necessary that the values of b_i obey

$$\sum_i b_i = 0; \tag{4.38}$$

otherwise, the Madelung potentials are not self-consistent with the site charges at the minimum.[4]

Proof: As established earlier, at the minimum in E, Eqn. (4.24) holds. Substituting Eqn. (4.25) into this gives

$$V_i = -aQ_i + ab_i. \tag{4.39}$$

Summing over all i, we find that

$$\sum_i V_i = -a\sum_i Q_i + a\sum_i b_i. \tag{4.40}$$

We will use this equation in a moment.

Now, if Eqn. (2.68) holds, then so also does the equation

$$\sum_i V_i = 0 \tag{4.41}$$

if the positions of the sites form an infinite Bravais lattice. The reasoning behind this is as follows. From Eqn. (2.54), it can be seen that the contribution to the Madelung potential of site i due to Q_j is $M_{ij}Q_j$. Therefore the total amount of Madelung potential 'allocated' to the system from site j is

$$\sum_i M_{ij}Q_j = Q_j\sum_i M_{ij} = Q_j\sum_{j\neq i}\frac{1}{|\mathbf{R}_j - \mathbf{R}_i|}, \tag{4.42}$$

where we have used Eqn. (2.55). This is equivalent to

$$Q_j\sum_{\mathbf{R}\neq 0}\frac{1}{|\mathbf{R}|} \tag{4.43}$$

[4]There is also a condition which must be met for self-consistency at the minimum in E for the CEFM. This condition is more complicated than Eqn. (4.38), and involves not only the values of b_i, but also the values of a_i and the underlying lattice.

if the positions of the sites form an infinite Bravais lattice whose lattice vectors are the set $\{\mathbf{R}\}$. Note that the summation in the above is independent of j. Denoting this system-dependent quantity as ζ, we have therefore just shown that the total amount of Madelung potential allocated to the system due to Q_j is $Q_j\zeta$. With this in mind, it can be seen that, if the system is charge neutral, then the total amount of Madelung potential allocated to the system from all sites is

$$\sum_j Q_j\zeta = 0. \tag{4.44}$$

Since the total amount of Madelung potential allocated to the system from all sites is equal to the sum of all Madelung potentials in the system, it therefore follows that Eqn. (4.41) must hold.

Substituting Eqns. (2.68) and (4.41) into Eqn. (4.40) gives

$$0 = 0 + a\sum_i b_i. \tag{4.45}$$

Ignoring the trivial case of $a = 0$, it can be seen that this is satisfied only if Eqn. (4.38) holds. In other words, if Eqn. (4.38) does *not* hold, then we have a contradiction: the values of V_i and the values of Q_i are not self-consistent. ■

The results pertaining to the GLCM given in Section 4.2 are therefore only valid if Eqn. (4.38) holds.[5] One noteworthy situation in which this is the case is when the effective medium approach is used to achieve the local approximation - as described in Section 4.1.2 - and the effective medium is constructed using the SSCPA. Here, $F_i(L)$ becomes $E_{L,i}$ for a site belonging to the same species as i which is embedded in the SSCPA medium and constrained to contain L electrons. Recall that b_i is the charge which occurs at the minimum in $F_i(L)$. If site i belongs to species X, then $b_i = b_X$ is therefore the charge of an X site embedded in the SSCPA medium, with no constraint upon its number of constituent electrons, i.e. b_X is equal to Q_X^{SSCPA} - the charge associated with species X obtained from a SSCPA calculation (see Section 2.2.1). In Ref. [58] it was pointed out that

$$\sum_X Q_X^{\text{SSCPA}} c_X = 0. \tag{4.46}$$

[5]Similarly, the results pertaining to the CEFM in Section 2.2.3.4 are only valid if the analogous equation to Eqn. (4.38) for the CEFM holds. Note that - as was mentioned in a previous footnote - this equation involves not only the values of b_i, but also the values of a_i and the underlying lattice.

Therefore

$$\sum_X b_X c_X = 0 \tag{4.47}$$

- which is the analogous equation to Eqn. (4.38) for the values of b_X - holds in this situation, and hence so also must Eqn. (4.38). This situation notwithstanding, there is no *a priori* reason for the values of b_i to obey Eqn. (4.38) in general. Nevertheless, we always assume when considering the GLCM that they *do* obey this equation. This is done in order to mirror the approach of Ref. [32], in which many useful results pertaining to the CEFM are given; doing this enables us to readily apply these results to the GLCM - something which will be done many times throughout this chapter. If one wishes to consider a set of b_i which do not obey Eqn. (4.38), then one must explicitly constrain the system to be charge neutral when minimising E. Incidently, all results pertaining to the GLCM in this thesis can be generalised to valid for such a set of b_i by making the following transformation:

$$b_i \rightarrow b_i', \tag{4.48}$$

where

$$b_i' = b_i - \langle b \rangle, \tag{4.49}$$

and $\langle b \rangle$ denotes the mean value of b_i for all i.

Proof: As mentioned above, if Eqn. (4.38) does not hold then one must explicitly impose the constraint of Eqn. (2.68) when minimising E, otherwise the values of Q_i and V_i are not self-consistent at the minimum in E. Consider the minimum in E subject to the constraint of Eqn. (2.68). This minimum can be found by using the method of Lagrange multipliers: it occurs when

$$\tilde{E} = E - \xi \sum_i Q_i \tag{4.50}$$

is minimised with respect to the variables Q_i for all i and ξ. In the case of the CEFM, minimising $\tilde{E}$ in the aforementioned manner gives a set of $N + 1$ equations, where N is the total number of sites in the system. These equations are Eqn. (2.68) and

$$V_i = -a_i Q_i + a_i b_i + \xi \tag{4.51}$$

for all i [64]. The analogous equations for the GLCM can be obtained by setting $a_i = a$ for all i. These are Eqn. (2.68) and

$$V_i = -a Q_i + a b_i + \xi \tag{4.52}$$

for all i. Summing over all i in the above equaion gives

$$\sum_i V_i = -a \sum_i Q_i + a \sum_i b_i + N\xi. \tag{4.53}$$

As was shown earlier, Eqn. (4.41) follows from Eqn. (2.68). With this in mind, Eqn. (2.68) implies that the above equation is equivalent to

$$0 = 0 + a \sum_i b_i + N\xi. \tag{4.54}$$

This becomes

$$\xi = -a \frac{1}{N} \sum_i b_i = -a\langle b \rangle \tag{4.55}$$

after rearranging for ξ, where we have used the definition of $\langle b \rangle$. Finally, substituting the above into Eqn. (4.52) and factorising the resulting equation gives

$$V_i = -aQ_i + a(b_i - \langle b \rangle). \tag{4.56}$$

The above equation describes the values of Q_i and V_i at the minimum in $\tilde{E}$, i.e. the minimum in E subject to the constraint of charge neutrality. The analogous equation for the minimum in E *without* the constraint of charge neutrality is

$$V_i = -aQ_i + ab_i, \tag{4.57}$$

which follows from substituting Eqn. (4.25) into Eqn. (4.24). As can be seen from comparing the above two equations, the latter becomes equivalent to the former if one uses $b_i' = b_i - \langle b \rangle$ in place of b_i. Therefore the minimum in E subject to the constraint of charge neutrality is equivalent to the minimum in E without the constraint of charge neutrality if in the latter the values of b_i are transformed according to Eqns. (4.48) and (4.49). Since it is necessary to impose the constraint of charge neutrality if Eqn. (4.38) does not hold, and all results pertaining to the GLCM in this thesis are obtained ignoring the constraint of charge neutrality, and hence rely on the validity of Eqn. (4.38); then the results in this thesis can be generalised to apply when Eqn. (4.38) does not hold by using the aforementioned transformation. ■

4.3 Accuracy of the Model

In a moment we will examine the properties of the GLCM in detail. However, before doing this, it is useful to understand how accurate a description of disordered alloys the

System	c	a_{Cu}	$a_{\mathrm{Zn/Pd}}$	$(b_{\mathrm{Cu}} - b_{\mathrm{Zn/Pd}})$
fcc $\mathrm{Cu}_c\mathrm{Pd}_{1-c}$	0.1	0.600	0.593	0.238
	0.25	0.597	0.588	0.229
	0.5	0.593	0.588	0.219
	0.75	0.597	0.586	0.212
	0.9	0.598	0.586	0.211
bcc $\mathrm{Cu}_c\mathrm{Zn}_{1-c}$	0.1	0.933	0.861	0.155
	0.25	0.951	0.903	0.159
	0.5	0.917	0.911	0.156
	0.75	0.882	0.887	0.155
	0.9	0.859	0.867	0.158
fcc $\mathrm{Cu}_c\mathrm{Zn}_{1-c}$	0.1	0.870	0.796	0.145
	0.25	0.862	0.809	0.150
	0.5	0.833	0.804	0.151
	0.75	0.813	0.791	0.150
	0.9	0.812	0.794	0.152

Table 4.1: CEFM parameters derived from the *ab initio* calculations of Refs. [22] and [23]. The values here were calculated from those given in Table 1 of Ref. [58].

GLCM can provide. As was mentioned in Section 2.2.3.4, the CEFM can provide an extremely accurate description of disordered alloys. Its success is due to the fact that the Q-V relations, which are implicit in the CEFM, hold to a high degree of accuracy in the results of *ab initio* calculations. Given that the GLCM differs from the CEFM only in that the values of a_X are assumed to be the same for all species, we therefore expect the accuracy of the GLCM to be limited by the validity of this assumption. The values of a_X obtained from the *ab initio* calculations of Refs. [22] and [23] are shown in Table 4.1. These results provide a test of the aforementioned assumption. As can be seen from the table, the values of a_{Cu} and $a_{\mathrm{Zn/Pd}}$ are, in general, very similar for a given system. In fact, the differences between a_{Cu} and $a_{\mathrm{Zn/Pd}}$ range from 1 to 10%. For systems at the lower end of this range, it is reasonable to expect that the GLCM will give quantitatively accurate predictions. While this is perhaps not the case for systems at the upper end of this range, we still expect the GLCM to be a useful tool for predicting qualitative trends in these systems. We should point out that, while one could use the CEFM to gain at least as accurate results for any particular system as the GLCM, the GLCM has the advantage that it is significantly simpler, as we will see throughout this chapter.

4.4 Expressions for Q_i

We will now derive explicit expressions for Q_i which apply in the GLCM. To do this, we will exploit an expression for Q_i derived in Ref. [32], which pertains to the CEFM. Setting $a_i = a$ for all i in this expression gives

$$Q_i = \sum_j G_{ij} a b_j, \tag{4.58}$$

where[6]

$$G = H^{-1}. \tag{4.59}$$

For the sake of completeness, we include a derivation of this result here.

Proof: We will now derive Eqn. (4.58). Consider Eqn. (4.24). Using Eqns. (2.54) and (4.25), this becomes

$$\sum_j M_{ij} Q_j = -a Q_i + a b_i. \tag{4.60}$$

Rearranging this, and exploiting the properties of the Kronecker delta, this in turn becomes

$$\sum_j M_{ij} Q_j + \sum_j a \delta_{ij} Q_j = a b_i, \tag{4.61}$$

which can be seen to be equivalent to

$$\sum_j H_{ij} Q_j = a b_i \tag{4.62}$$

from inspection of Eqn. (4.28). Premultiplying the above equation by G, and noting that G is the inverse of H (Eqn. (4.59)), gives Eqn. (4.58). ■

To proceed further, we will exploit two properties of G. The first of these is

$$\sum_j G_{ij} = 0, \tag{4.63}$$

which is derived in Ref. [32]. This follows from the properties of the Madelung matrix, and ensures that Eqn. (2.68) always holds [32]. The second property of G which we will require is:

Theorem 4.1. *The elements of G are such that*

$$G_{ij} = G_\beta \quad \text{if } j \in \beta_i \tag{4.64}$$

[6]Note that the existence of G is guaranteed by the fact that H is constrained to be positive-definite, since all positive-definite matrices are invertible.

for some set of values $\{G_0, G_1, G_2, \dots\}$.

This theorem simply states that there is a high degeneracy in the elements of G for a given system: G_{ij} is the same for all sites j at the same distance from i.

Proof: We will now prove Theorem 4.1. Consider the H matrix. Recall that H is positive-definite. Since all positive-definite matrices are invertible, H is therefore invertible. Its inverse G can therefore be expressed as a power series as follows [65]:

$$G = H^{-1} = \sum_{n=0}^{\infty} (I - H)^n = \sum_{n=0}^{\infty} J^n, \tag{4.65}$$

where we have defined

$$J = I - H, \tag{4.66}$$

and I is the identity matrix. We will show by induction that each term in the series individually obeys the property which we are to prove that G has; i.e. we will show that, for all $n \geq 0$, J^n is such that

$$(J^n)_{ij} = J_\beta^{(n)} \quad \text{if } j \in \beta_i \tag{4.67}$$

for some set of values $\{J_0^{(n)}, J_1^{(n)}, J_2^{(n)}, \dots\}$. If this is the case then, for $j \in \beta_i$,

$$G_{ij} = \sum_{n=0}^{\infty} (J^n)_{ij} = \sum_{n=0}^{\infty} J_\beta^{(n)}, \tag{4.68}$$

which implies that Theorem 4.1 is true with

$$G_\beta = \sum_{n=0}^{\infty} J_\beta^{(n)}. \tag{4.69}$$

Consider first $n = 0$. Note that J^0 is simply the identity matrix I. $I_{ij} = 0$ unless $j = i$, or in other words, if $j \in 0_i$. Hence we see that Eqn. (4.67) is true for $n = 0$, with

$$J_\beta^{(0)} = \begin{cases} 1 & \text{if } \beta = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.70}$$

Consider now $n = 1$. Using the definitions of H (Eqn. (4.28)) and J (Eqn. (4.66)) we find that

$$J_{ij} = \begin{cases} 1 - a & \text{if } j = i \\ -1/|\mathbf{R}_i - \mathbf{R}_j| & \text{otherwise.} \end{cases} \tag{4.71}$$

From this it can be seen that, since $|\mathbf{R}_i - \mathbf{R}_j| = R_\beta$ if $j \in \beta_i$, Eqn. (4.67) is true for $n = 1$ with

$$J_\beta^{(1)} = \begin{cases} 1 - a & \text{if } \beta = 0 \\ -1/R_\beta & \text{otherwise} \end{cases} \tag{4.72}$$

Consider now $n = 2$. The elements of J^2 are given by

$$(J^2)_{ij} = \sum_k J_{ik} J_{kj}. \tag{4.73}$$

We can split the sum over k into contributions from sites in different combinations of shells of i and j as follows:

$$(J^2)_{ij} = \sum_{\gamma=0}^{\infty} \sum_{\delta=0}^{\infty} \sum_{k \in \gamma_i \cap \delta_j} J_{ik} J_{kj}. \tag{4.74}$$

Exploiting the fact that, as was shown above, Eqn. (4.67) is true for $n = 1$, the above equation becomes

$$(J^2)_{ij} = \sum_{\gamma=0}^{\infty} \sum_{\delta=0}^{\infty} J_\gamma^{(1)} J_\delta^{(1)} \sum_{k \in \gamma_i \cap \delta_j} 1. \tag{4.75}$$

Without loss of generality, let i and j be separated R_β. Now,

$$K_\beta^\delta(\gamma) = \sum_{k \in \gamma_i \cap \delta_j} 1 \quad \text{if } j \in \beta_i \text{ (or equivalently } i \in \beta_j), \tag{4.76}$$

which follows from the definition of $K_\beta^\delta(\gamma)$: $K_\beta^\delta(\gamma)$ is the number of sites in shell δ of a particular site j which are also in shell γ of another site i, where sites i and j are separated by R_β. With this in mind, Eqn. (4.75) becomes

$$(J^2)_{ij} = \sum_{\gamma=0}^{\infty} \sum_{\delta=0}^{\infty} J_\gamma^{(1)} J_\delta^{(1)} K_\beta^\delta(\gamma) \quad \text{if } j \in \beta_i. \tag{4.77}$$

From this we see that Eqn. (4.67) is true for $n = 2$, with

$$J_\beta^{(2)} = \sum_{\gamma=0}^{\infty} \sum_{\delta=0}^{\infty} J_\gamma^{(1)} J_\delta^{(1)} K_\beta^\delta(\gamma). \tag{4.78}$$

Consider now $n = 3$. The elements of J^3 are given by

$$(J^3)_{ij} = \sum_k (J^2)_{ik} J_{kj}. \tag{4.79}$$

Following a similar procedure as for $n = 2$, it can be shown that

$$(J^3)_{ij} = \sum_{\gamma=0}^{\infty} \sum_{\delta=0}^{\infty} J_{\gamma}^{(2)} J_{\delta}^{(1)} K_{\beta}^{\delta}(\gamma), \tag{4.80}$$

where we have exploited the truth of Eqn. (4.67) for $n = 2$, and again assumed that i and j are separated by R_β. From the above equation we see that Eqn. (4.67) is true for $n = 3$, with

$$J_{\beta}^{(3)} = \sum_{\gamma=0}^{\infty} \sum_{\delta=0}^{\infty} J_{\gamma}^{(2)} J_{\delta}^{(1)} K_{\beta}^{\delta}(\gamma). \tag{4.81}$$

We can continue in this manner *ad infinitum* to prove that Eqn. (4.67) is also true for all $n > 3$. Hence Eqn. (4.67) is true for all $n \geq 0$ and so, for the reasons discussed earlier, Theorem 4.1 must be true. ■

Using Eqn. (4.63) and Theorem 4.1, we find that

$$Q_i = aG_0 \sum_{Y} (b_Y - b_i) \sum_{\beta=1}^{\infty} g_\beta N_{iY\beta}, \tag{4.82}$$

where we have defined g_β as

$$g_\beta = \frac{G_\beta}{G_0}. \tag{4.83}$$

Proof: We will now derive Eqn. (4.82). To do this, we require the following equation:

$$\sum_{\beta=1}^{\infty} g_\beta Z_\beta = -1. \tag{4.84}$$

This is derived as follows. Consider Eqn. (4.63). This implies that

$$G_{ii} = -\sum_{j \neq i} G_{ij}. \tag{4.85}$$

Splitting the summation in the above into contributions from each shell of i gives

$$G_{ii} = -\sum_{\beta=1}^{\infty} \sum_{j \in \beta_i} G_{ij}, \tag{4.86}$$

which becomes

$$G_0 = -\sum_{\beta=1}^{\infty} G_\beta \sum_{j \in \beta_i} 1 \tag{4.87}$$

after applying Theorem 4.1. This in turn becomes

$$\sum_{\beta=1}^{\infty} G_\beta Z_\beta = -G_0 \tag{4.88}$$

after using the fact that

$$Z_\beta = \sum_{j\in\beta_i} 1 \quad \text{for all } i, \tag{4.89}$$

which is a restatement of the definition of Z_β: Z_β is the number of sites in shell β of any site i. Dividing Eqn. (4.88) by G_0 and using Eqn. (4.83) gives Eqn. (4.84).

Now, to derive Eqn. (4.82), we firstly write Eqn. (4.58) as

$$Q_i = G_{ii} a b_i + \sum_{j\neq i} G_{ij} a b_j. \tag{4.90}$$

Splitting the summation over j into contributions from each shell of i gives

$$Q_i = G_{ii} a b_i + \sum_{\beta=1}^{\infty} \sum_{j\in\beta_i} G_{ij} a b_j, \tag{4.91}$$

which becomes

$$Q_i = G_0 a b_i + \sum_{\beta=1}^{\infty} \sum_{j\in\beta_i} G_\beta a b_j. \tag{4.92}$$

after applying Theorem 4.1. Using Eqn. (4.83) in the second term and Eqn. (4.84) in the first term gives

$$\begin{aligned} Q_i &= -G_0 a b_i \sum_{\beta=1}^{\infty} g_\beta Z_\beta + \sum_{\beta=1}^{\infty} \sum_{j\in\beta_i} a G_0 g_\beta b_j \\ &= -a G_0 b_i \sum_{\beta=1}^{\infty} g_\beta \sum_{j\in\beta_i} 1 + a G_0 \sum_{\beta=1}^{\infty} g_\beta \sum_{j\in\beta_i} b_j \\ &= -a G_0 \sum_{\beta=1}^{\infty} g_\beta \sum_{j\in\beta_i} b_i + a G_0 \sum_{\beta=1}^{\infty} g_\beta \sum_{j\in\beta_i} b_j \\ &= a G_0 \sum_{\beta=1}^{\infty} g_\beta \sum_{j\in\beta_i} (b_j - b_i) \end{aligned} \tag{4.93}$$

where in the second line we have used Eqn. (4.89). Splitting the summation over j in

the above equation into contributions from each species gives

$$\begin{aligned} Q_i &= aG_0 \sum_{\beta=1}^{\infty} g_\beta \sum_Y \sum_{j \in \beta_i \cap Y} (b_Y - b_i) \\ &= aG_0 \sum_Y (b_Y - b_i) \sum_{\beta=1}^{\infty} g_\beta \sum_{j \in \beta_i \cap Y} 1. \end{aligned} \tag{4.94}$$

Noting that, by definition,

$$N_{iY\beta} = \sum_{j \in \beta_i \cap Y} 1, \tag{4.95}$$

the above expression for Q_i can be seen to be equivalent to Eqn. (4.82). ■

Setting $b_i = b_X$ in Eqn. (4.82) gives an expression for Q_i which applies only for X sites. This is

$$Q_i = aG_0 \sum_Y b_{YX} \sum_{\beta=1}^{\infty} g_\beta N_{iY\beta}, \tag{4.96}$$

where b_{YX} is defined by the equation

$$b_{YX} = b_Y - b_X. \tag{4.97}$$

Eqn. (4.96) allows us to interpret the charge distribution throughout the system under consideration in terms of charge transfer between all pairs of sites. Consider an X site. From Eqn. (4.96), the charge of the X site can be thought of as resulting from 'donations' of charge from all other sites as follows: each Y site at distance R_β donates a quantity of charge

$$t_\beta^{Y \to X} = aG_0 b_{YX} g_\beta \tag{4.98}$$

to the X site. Conversely, the X site itself donates a quantity of charge $t_\beta^{X \to Y} = aG_0 b_{XY} g_\beta$ to each Y site at distance R_β. Since $b_{XY} = -b_{YX}$, the charge donated to the X site from the Y site, and the charge donated to the Y site from the X site, are equal and opposite, i.e. $t_\beta^{Y \to X} = -t_\beta^{X \to Y}$. Therefore a quantity of charge $|t_\beta^{Y \to X}| = |t_\beta^{X \to Y}|$ can be regarded as being transferred directly between each pair of X and Y sites separated by R_β. Interestingly, as in the OLCM, charge is transferred only between pairs of *unlike* sites, i.e. no charge is transferred between pairs of sites which belong to the same species. This follows from the fact that, since $b_{XX} = 0$, $t_\beta^{X \to X} = 0$. More generally, the species-dependence of the amount of charge transferred between an X and Y site enters entirely through their *electropositivity difference* b_{XY}: the higher their electropositivity difference; the higher the amount of charge transferred.

In addition to the charge transfer interpretation discussed above, the charge distribution described by Eqn. (4.96) can be expressed as a superposition of *localised charge distributions* (LCDs) associated with each site - as is the case in the MLCM (see Section 3.1). Using the same notation as in Section 3.1, the LCD of an X site in the GLCM consists of a charge

$$Q_{X0} = b_X(aG_0) \tag{4.99}$$

on the X site itself, and a charge

$$Q_{X\beta} = b_X(aG_0)g_\beta \tag{4.100}$$

on each site in shell $\beta \geq 1$ of the X site.

Proof: We will now show that the values of Q_i in the GLCM can be expressed as a superposition of LCDs as described above. Consider Eqn. (4.58). We can write this as

$$Q_i = \sum_j Q_i^j, \tag{4.101}$$

where

$$Q_i^j = b_j aG_{ij} \tag{4.102}$$

can be considered as the LCD associated with site j. Assuming that j belongs to species X, the above equation becomes

$$Q_i^j = b_X aG_{ij}. \tag{4.103}$$

Applying Theorem 4.1 to this gives

$$Q_i^j = Q_{X0} = b_X aG_0 \quad \text{if } i = j \tag{4.104}$$

and

$$Q_i^j = Q_{X\beta} = b_X aG_\beta \quad \text{if } i \in \beta_j \text{ and } \beta \geq 1. \tag{4.105}$$

The former of the above two equations is equivalent to Eqn. (4.99). The latter becomes Eqn. (4.100) after using Eqn. (4.83). ■

Similarly to the MLCM, Q_{X0} is the mean charge of an X site in a random alloy with the same underlying lattice and free parameters, i.e. a and b_X for all X, as the alloy under consideration. This will be shown in Section 4.8.

4.5 Relationship with the OLCM

The fact that the values of Q_i in the GLCM can be understood in terms of charge transfer between pairs of unlike sites, and as a superposition of species-dependent LCDs, is reminiscent of the MLCM and the OLCM (where recall that the latter is a particular case of the former). It can be shown that, for binary alloys, the GLCM is equivalent to the OLCM with $\beta_{\max} = \infty$, with the following equation relating the variables associated with each of the two models:

$$\lambda_\beta = \frac{1}{2} a G_0 b_{AB} g_\beta. \tag{4.106}$$

Proof: In the particular case of a binary alloy, using Eqn. (4.96), the charges on A and B sites within the GLCM obey the equations

$$Q_i = a G_0 b_{BA} \sum_{\beta=1}^{\infty} g_\beta N_{iB\beta} \tag{4.107}$$

and

$$Q_i = a G_0 b_{AB} \sum_{\beta=1}^{\infty} g_\beta N_{iA\beta} \tag{4.108}$$

respectively. Reverting to the notation of the previous chapter, the above two equations can be expressed concisely as

$$Q_i = a G_0 S_i b_{AB} \sum_{\beta=1}^{\infty} g_\beta N_{i\beta}, \tag{4.109}$$

where $S_i = -1$ for an A site and $+1$ for a B site, and we have used the fact that

$$b_{YX} = -b_{XY} \tag{4.110}$$

(see Eqn. (4.97)). Comparing the above expression for Q_i to Eqn. (3.131), i.e. the OLCM charge law for $\beta_{\max} = \infty$, it can be seen to be equivalent to Eqn. (3.131) with λ_β given by Eqn. (4.106).

All that remains to show is that the vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots)$ obeys Eqn. (3.37), i.e. that it is an eigenvector of the infinite f-matrix with eigenvalue a. Recall that this is a requirement in the OLCM. Note that if $(G_1, G_2, \dots)$ is an eigenvector of the infinite f-matrix then so is $\boldsymbol{\lambda}$. This follows from the fact that, as can be seen from Eqns. (4.106) and (4.83), they differ only by a factor of $ab_{AB}/2$. Furthermore, they will have the same eigenvalue. It is therefore sufficient to show that $(G_1, G_2, \dots)$ is an eigenvector of the infinite f-matrix with eigenvalue a, which we will now do. Since G

is the inverse of H,

$$\sum_j H_{ij}G_{jk} = \delta_{ik}. \tag{4.111}$$

Consider the case where i and k are separated by R_β, where $\beta > 0$, i.e. $k \neq i$. In this case the above becomes

$$H_{ii}G_{ik} + H_{ik}G_{kk} + \sum_{j\neq i,k} H_{ij}G_{jk} = 0 \tag{4.112}$$

after separating out the $j = i$ and $j = k$ terms from the summation. Applying Eqn. (4.28) to the first term, and Eqn. (4.85) to the second term, we find that

$$aG_{ik} - \sum_{j\neq k} H_{ik}G_{jk} + \sum_{j\neq i,k} H_{ij}G_{jk} = 0. \tag{4.113}$$

Rearranging this gives

$$\sum_{j\neq k} H_{ik}G_{jk} - \sum_{j\neq i,k} H_{ij}G_{jk} = aG_{ik}, \tag{4.114}$$

which becomes

$$\sum_{\gamma=1}^{\infty}\sum_{j\in\gamma_k} H_{ik}G_{jk} - \sum_{\gamma=1}^{\infty}\sum_{\delta=1}^{\infty}\sum_{j\in\gamma_k\cap\delta_i} H_{ij}G_{jk} = aG_{ik} \tag{4.115}$$

after splitting the summations into contributions from shells of k and i. Now, from the definition of H (see Eqns. (4.28) and (2.55)), it follows that: $H_{ik} = 1/R_\beta$, since we assumed earlier that i and k are separated by R_β; and $H_{ij} = 1/R_\delta$ if $j \in \delta_i$. Substituting these results into the above equation, and applying Theorem 4.1, we find that

$$\begin{aligned}
\sum_{\gamma=1}^{\infty}\sum_{j\in\gamma_k}\frac{1}{R_\beta}G_\gamma - \sum_{\gamma=1}^{\infty}\sum_{\delta=1}^{\infty}\sum_{j\in\gamma_k\cap\delta_i}\frac{1}{R_\delta}G_\gamma &= aG_\beta \\
\sum_{\gamma=1}^{\infty}\frac{1}{R_\beta}G_\gamma\sum_{j\in\gamma_k}1 - \sum_{\gamma=1}^{\infty}\sum_{\delta=1}^{\infty}\frac{1}{R_\delta}G_\gamma\sum_{j\in\gamma_k\cap\delta_i}1 &= aG_\beta \\
\sum_{\gamma=1}^{\infty}\frac{Z_\gamma}{R_\beta}G_\gamma - \sum_{\gamma=1}^{\infty}\sum_{\delta=1}^{\infty}\frac{K_\beta^\delta(\gamma)}{R_\delta}G_\gamma &= aG_\beta \\
\sum_{\gamma=1}^{\infty}\left[\frac{Z_\gamma}{R_\beta} - \sum_{\delta=1}^{\infty}\frac{K_\beta^\delta(\gamma)}{R_\delta}\right]G_\gamma &= aG_\beta,
\end{aligned} \tag{4.116}$$

where in the second line we have applied Eqn. (4.89) to the first term and Eqn. (4.76)

to the second term. Substituting Eqn. (3.22) (in conjunction with Eqn. (3.23)) into the expression in the square brackets gives

$$\sum_{\gamma=1}^{\infty} f_\beta(\gamma) G_\gamma = a G_\beta, \tag{4.117}$$

where recall that $\beta > 0$. Therefore $(G_1, G_2, \dots)$, and hence $(\lambda_1, \lambda_2, \dots)$, is an eigenvector of the infinite f-matrix with eigenvalue a.

We conclude this proof by justifying our assertion made in Section 3.5 that the eigenvalues of the infinite f-matrix form a continuum. Recall that the elements of the vector $(G_1, G_2, \dots)$ are simply elements of the matrix G (see Theorem 4.1). Recall also that we have constrained a to be such that H is positive-definite, and that this amounts to a obeying Eqn. (4.29). Now, all positive-definite matrices are invertible. The aforementioned constraint on a therefore guarantees the existence of G - which is the inverse of the matrix H (Eqn. (4.59)) - and hence also guarantees the existence of the vector $(G_1, G_2, \dots)$. Thus for each possible a which obeys Eqn. (4.29) there will be a vector $(G_1, G_2, \dots)$ which, as shown above, is necessarily an eigenvector of the infinite f-matrix with eigenvalue a. In other words, each possible a which obeys Eqn. (4.29) is an eigenvalue of the infinite f-matrix. Since Eqn. (4.29) describes a continuum of possible a, it is therefore the case that the eigenvalues of the infinite f-matrix form a continuum above $a_{\min}$.[7] ■

The above result allows us to make the following three points regarding the OLCM with $\beta_{\max} = \infty$. Firstly, given that this model can only address binary alloys, the GLCM can therefore be considered to be its generalisation which can treat alloys containing any number of species. Secondly, the values of Q_i in the OLCM with $\beta_{\max} = \infty$ are those which (for binary alloys) minimise an energy function of the form given in Eqns. (4.21), (4.22) and (4.23). Thirdly, the assumptions which underpin the OLCM with $\beta_{\max} = \infty$ are the same as those described in Sections 4.1 and 4.2 which underpin the GLCM.

4.6 Screening

Henceforth we will assume that a and the values of b_X are known for the system under consideration. These could have been extracted from the results of *ab initio* calculations, or determined by some other means. With this information, however, we

[7] Note that the preceding discussion does not preclude the possibility that the infinite f-matrix has eigenvalues which do not obey Eqn. (4.29): we have shown that the eigenvalues of the infinite f-matrix *at least* consist of the continuum described by Eqn. (4.29).

still do not know the values of g_β and G_0 which appear in the above expressions for Q_i. Knowledge of these values is required before the above expressions for Q_i, as well as those derived later in Sections 4.7 and 4.8, can be used in practice. In this section, we will calculate the values of g_β and G_0 for a wide range of systems. In doing this, we will learn much about the nature of the screening in the GLCM.

The values of g_β and G_0 are completely determined by a and the underlying lattice. The reasoning behind this is as follows. As can be seen from Eqn. (4.83) and Theorem 4.1, G_0 and g_β depend entirely on the matrix G. Recall that G is the inverse of H (Eqn. (4.59)), and that H depends entirely on a and the Madelung matrix M (Eqn. (4.28)). Therefore G_0 and g_β depend entirely on a and M. Now, as can be seen from the definition of the elements of M given in Eqn. (2.55), M depends entirely on the underlying lattice. Therefore G_0 and g_β depend entirely on a and the underlying lattice. We emphasise that G_0 and g_β do not depend on the values of b_i, i.e. they do not depend on the particular species of each site, and hence all systems which have the same a and underlying lattice will have the same values of G_0 and g_β. This is *not* the case for the analogous variables in the CEFM. In fact, one could consider this to be a strength of the GLCM: the parameters G_0 and g_β are transferable to all systems with the same value of a and underlying lattice; one need only calculate the values of G_0 and g_β once for each possible value of a and underlying lattice. However, it is a daunting task to calculate G_0 and g_β for all combinations of a, lattice type, and unit cell volume - which we will characterise here by the lattice's Wigner-Sietz radius R_{WS}. We can reduce the parameter space by appealing to the following result:

Theorem 4.2. *Systems with the same lattice type and value of aR_{WS} have the same vector*

$$\mathbf{g} = (aG_0, g_1, g_2, \dots). \tag{4.118}$$

Proof: We will now prove Theorem 4.2. Using Eqn. (4.28) we can express the matrix H as

$$H = aI + M = aI + \frac{1}{R_{\mathrm{WS}}} M^{R_{\mathrm{WS}}=1}, \tag{4.119}$$

where $M^{R_{\mathrm{WS}}=1}$ is the Madelung matrix for the same lattice type as the system under consideration, but with $R_{\mathrm{WS}} = 1$ (M and $M^{R_{\mathrm{WS}}=1}$ are related by $M = M^{R_{\mathrm{WS}}=1}/R_{\mathrm{WS}}$).

By definition (Eqn. (4.59)) G is the inverse of H, which means that

$$\begin{aligned} G &= \left(aI + \frac{1}{R_{\mathrm{WS}}} M^{R_{\mathrm{WS}}=1} \right)^{-1} \\ &= \left[a \left(I + \frac{1}{(aR_{\mathrm{WS}})} M^{R_{\mathrm{WS}}=1} \right) \right]^{-1} \\ &= \frac{1}{a} \left(I + \frac{1}{(aR_{\mathrm{WS}})} M^{R_{\mathrm{WS}}=1} \right)^{-1}. \end{aligned} \tag{4.120}$$

From this it follows that

$$aG = \left(I + \frac{1}{(aR_{\mathrm{WS}})} M^{R_{\mathrm{WS}}=1} \right)^{-1}. \tag{4.121}$$

Recall that G_β are elements of G. From the above equation it can be seen that aG depends only on the underlying lattice *type* (through the matrix $M^{R_{\mathrm{WS}}=1}$, which is the same for all lattices of the same type), and the value of (aR_{WS}). Therefore the following lemma holds: *the values of aG_β for all β are the same for all systems with the same lattice type and value of (aR_{WS}).*

Setting $\beta = 0$ in the above lemma reveals that aG_0, which is the first element in $\mathbf{g}$ described above, is the same for all systems with the same lattice type and value of aR_{WS}. For the remaining elements in $\mathbf{g}$, consider the quantity $(aG_\beta)/(aG_0)$. From the lemma, both the numerator and the denominator, and hence the quantity as a whole, are the same for all systems with the same lattice type and value of aR_{WS}. Noting that

$$\frac{aG_\beta}{aG_0} = \frac{G_\beta}{G_0} = g_\beta, \tag{4.122}$$

where we have used Eqn. (4.83), it is therefore the case that g_β for $\beta > 0$, i.e. the remaining elements in $\mathbf{g}$, are the same for all systems with the same lattice type and value of aR_{WS}. ■

Theorem 4.2 implies that we need only calculate $\mathbf{g}$ once for each combination of aR_{WS} and lattice type. This is done below, though we restrict ourselves to the fcc, bcc and sc lattices. It should be pointed out that, while aG_0 appears in $\mathbf{g}$ instead of G_0, the latter can easily be calculated from the former if required since a will be known for the particular system in question. However, in the analytical expressions derived in this chapter, G_0 only ever appears in a product with a, and hence the value of G_0 itself is never required.

4.6.1 Physical Significance of aR_{WS}, aG_0 and g_β

Before calculating how $\mathbf{g}$ varies with aR_{WS}, it is instructive to consider the physical significance of these quantities. We will begin with aR_{WS}. Recall that a determines the strength of the local interactions which act to keep the site charges Q_i at their bare values b_i. The analogous quantity for the inter-site Coulomb interactions is $1/R_{\mathrm{WS}}$. The reasoning behind this is as follows. A high value of $1/R_{\mathrm{WS}}$ corresponds to a low value of R_{WS}, i.e. a lattice in which the sites are close together. Here, the inter-site Coulomb interactions will be strong. Conversely, a low value of $1/R_{\mathrm{WS}}$ corresponds to a high value of R_{WS}, i.e. a lattice in which the sites are further apart. Here, the inter-site Coulomb interactions will be weak. With the above in mind, it follows that $aR_{\mathrm{WS}} = a/(1/R_{\mathrm{WS}})$ is a dimensionless quantity which determines the strength of the local interactions *relative* to the strength of the inter-site Coulomb interactions. The higher the value of aR_{WS}, the more important the local interactions are, and the less important the inter-site Coulomb interactions are, in determining the values of Q_i which minimise E. Furthermore, in the limit $aR_{\mathrm{WS}} \to \infty$ the values of Q_i will be determined solely by the local interactions, while in the limit $aR_{\mathrm{WS}} \to 0$ the values of Q_i will be determined solely by the inter-site interactions. In the former limit we expect that all site charges will take their bare values, i.e. $Q_i = b_i$ for all i.[8]

We will now discuss the physical significance of the elements of $\mathbf{g}$. Consider the change in energy δE as a result of perturbing the charge on each site i by an amount δQ_i. Expanding E as a multi-dimensional Taylor series about its minimum, we find that

$$\delta E = \frac{1}{2}\sum_{i,j}\frac{\partial^2 E}{\partial Q_i \partial Q_j}\delta Q_i \delta Q_j + \frac{1}{6}\sum_{i,j,k}\frac{\partial^3 E}{\partial Q_i \partial Q_j \partial Q_k}\delta Q_i \delta Q_j \delta Q_k + \ldots, \tag{4.123}$$

where all partial derivatives are evaluated at the minimum, and we have used the fact that $\partial E/\partial Q_i = 0$ for all i. As can be seen from Eqns. (4.21), (4.22) and (4.23), E is quadratic in the charges, and hence all third and higher order partial derivatives of E vanish, leaving

$$\delta E = \frac{1}{2}\sum_{i,j}\frac{\partial^2 E}{\partial Q_i \partial Q_j}\delta Q_i \delta Q_j. \tag{4.124}$$

The matrix whose (i,j)th element is $\partial^2 E/\partial Q_i \partial Q_j$ is known as the *Hessian matrix* of the energy (evaluated at the energy minimum). In Ref. [32] it was pointed out that the Hessian matrix for the CEFM is equal to H defined by Eqn. (4.27). Therefore H

[8] The limit $aR_{\mathrm{WS}} \to \infty$ can be achieved by setting $E_M = 0$ in Eqn. (4.21). The resulting expression for E is $E = E_L$, which, as can be seen from Eqn. (4.22), is minimised when $Q_i = b_i$ for all i. Note that the system is still automatically charge neutral at this minimum on account of Eqn. (4.38).

defined by Eqn. (4.28) is the Hessian matrix for the GLCM.[9] The above equation can therefore be rewritten as

$$\delta E = \frac{1}{2}\sum_{i,j} H_{ij}\delta Q_i \delta Q_j. \tag{4.125}$$

Now, consider δE in the following two situations:

- Q_k is perturbed by an amount δQ_k, with the constraint that all other charges in the system are fixed, i.e. $\delta Q_i = 0$ for all $i \neq k$
- Q_k is perturbed by an amount δQ_k, with all other charges in the system being allowed to 'relax' as to keep the total energy of the system at a minimum.

In the latter case, the relaxing charges *screen* the perturbation δQ_k, while in the former case there is *no screening*. We will begin by considering what happens in the absence of screening. Setting $\delta Q_i = 0$ for all $i \neq k$ in Eqn. (4.125) gives

$$\delta E = \frac{1}{2} H_{kk}\delta Q_k^2. \tag{4.126}$$

This becomes

$$\delta E = \frac{1}{2} a \delta Q_k^2 \tag{4.127}$$

after using Eqn. (4.28). Consider now what happens if there *is* screening. In Ref. [32] it was shown that the perturbation δQ_k is screened such that

$$\delta Q_i = \frac{G_{ik}}{G_{kk}}\delta Q_k \tag{4.128}$$

for all i in the CEFM. Applying this result to the GLCM, it follows from Theorem 4.1 and Eqn. (4.83) that

$$\delta Q_j = g_\beta \delta Q_k \tag{4.129}$$

if $j \in \beta_k$ and $\beta > 1$. The above equation reveals the physical significance of g_β: for a perturbation δQ_k in the charge of site k, g_β is the amount of charge on a site in shell β of site k, normalised to the perturbation, which screens the perturbation. Hence g_β has exactly the same physical significance as the quantity ϕ_β described in previous chapters. This result could also have been derived by using the same procedure as was used in Section 3.1.1 to derive ϕ_β for the MLCM, but with LCDs pertaining to the GLCM instead of the OLCM; the LCD associated with an X site in the GLCM is described by Eqns. (4.99) and (4.100). Using Eqns. (4.125) and (4.128), an equation analogous to Eqn. (4.127) can be derived which applies in the presence of screening.

[9] This can be verified explicitly by using Eqns. (4.21), (4.22) and (4.23) to determine $\partial^2 E/\partial Q_i \partial Q_j$, and then comparing the result with Eqn. (4.28).

This is

$$\delta E = \frac{1}{2} a^{\text{scr}} \delta Q_k^2, \tag{4.130}$$

where

$$a^{\text{scr}} = \frac{1}{G_0} \tag{4.131}$$

Proof: We will now derive Eqn. (4.130). Substituting Eqn. (4.128) into Eqn. (4.125) gives an equation for δE which applies in the presence of screening:

$$\begin{aligned} \delta E =& \frac{1}{2} \sum_{i,j} H_{ij} \frac{G_{ik}}{G_{kk}} \delta Q_k \frac{G_{jk}}{G_{kk}} \delta Q_k \\ =& \frac{1}{2} \frac{1}{G_{kk}^2} \delta Q_k^2 \sum_{i,j} H_{ij} G_{ik} G_{jk} \\ =& \frac{1}{2} \frac{1}{G_{kk}^2} \delta Q_k^2 \sum_{i} G_{ik} \sum_{j} H_{ij} G_{jk}. \end{aligned} \tag{4.132}$$

Using Eqn. (4.111), the above becomes

$$\begin{aligned} \delta E =& \frac{1}{2} \frac{1}{G_{kk}^2} \delta Q_k^2 \sum_{i} G_{ik} \delta_{ik} \\ =& \frac{1}{2} \frac{1}{G_{kk}} \delta Q_k^2 \end{aligned} \tag{4.133}$$

Finally, using the fact that $G_{kk} = G_0$ - which follows from Theorem 4.1 - the above equation can be seen to be equivalent to Eqn. (4.130) (in conjunction with Eqn. (4.131)). ■

Comparing Eqns. (4.127) and (4.130), it can be seen that a determines the amount of energy required to perturb Q_k by δQ_k in the absence of screening, and a^{scr} determines the amount of energy required to perturb Q_k by δQ_k in the presence of screening. Thus, just as a is a measure of the strength of the 'bare' on-site interactions, so a^{scr} is a measure of the strength of the 'screened' on-site interactions.[10] From Eqn. (4.131), it can be seen that

$$aG_0 = \frac{a}{a^{\text{scr}}}. \tag{4.134}$$

This reveals that aG_0 is a measure of the *amount of screening* which occurs in the considered system. In the absence of screening, $a^{\text{scr}} = a$, and hence $aG_0 = 1$. In the presence of screening, it takes less energy to perturb Q_k than if there were no screening,

[10] In Ref. [32] it was pointed out that $a^{\text{scr}} = 1/G_{ii}$ in the CEFM, from which one could readily derive Eqn. (4.131). We have instead chosen to derive Eqn. (4.131) in a manner which makes the physical significance of the quantity a^{scr} more clear.

in which case $a^{\text{scr}} < a$, and hence $aG_0 > 1$. For perfect screening it takes no energy to perturb Q_k, in which case $a^{\text{scr}} = 0$, and hence $aG_0 = \infty$.

4.6.2 Asymptotic Expressions for aG_0 and g_β

In general, the elements of $\mathbf{g}$ must be calculated numerically. This will be done later. However, analytical expressions for aG_0 and g_β can be derived which are increasingly accurate for higher values of aR_{WS}, and exact in the limit $aR_{\text{WS}} \to \infty$. These are

$$aG_0 = \frac{(aR_{\text{WS}})(1+\nu)}{(aR_{\text{WS}})(1+\nu) - \nu} \tag{4.135}$$

and

$$g_\beta = -\left[\frac{\nu^2 e^\nu}{3(1+\nu)}\right] \frac{\exp\left(-\nu \frac{R_\beta}{R_{\text{WS}}}\right)}{\left(\frac{R_\beta}{R_{\text{WS}}}\right)}, \tag{4.136}$$

where

$$\nu = \sqrt{\frac{3}{(aR_{\text{WS}}) - 1.5}}. \tag{4.137}$$

Proof: We will now derive Eqns. (4.135) and (4.136). To do this, we will first find an analytical solution to the equation

$$V_i = -aQ_i + k_0\delta_{i0} \tag{4.138}$$

in the limit $aR_{\text{WS}} \to \infty$, where without loss of generality we have chosen site 0 to reside at the origin. We will then use this solution to derive expressions for aG_0 and g_β in the limit $aR_{\text{WS}} \to \infty$.

To find an analytical solution to Eqn. (4.138), a good starting point is its continuous analogue. This is

$$V(\mathbf{r}) = -A\rho(\mathbf{r}) + K\delta(\mathbf{r}), \tag{4.139}$$

where $\rho(\mathbf{r})$ is the charge density at position $\mathbf{r}$,

$$V(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} \tag{4.140}$$

is the conventional electrostatic potential at $\mathbf{r}$, and A and K are constants. Using the fact that $\rho(\mathbf{r})$ and $V(\mathbf{r})$ are related by Poisson's equation, i.e.

$$\nabla^2 V(\mathbf{r}) = -4\pi\rho(\mathbf{r}), \tag{4.141}$$

we can eliminate $\rho(\mathbf{r})$ from Eqn. (4.139). Doing this gives

$$(\nabla^2 - \mu^2)V(\mathbf{r}) = -4\pi\frac{K}{A}\delta(\mathbf{r}) \tag{4.142}$$

after rearranging, where

$$\mu^2 = 4\pi/A. \tag{4.143}$$

Eqn. (4.142) is a particular case of the inhomogeneous Helmholtz equation, whose solutions are well known. In our case, the solution is [66]

$$\begin{aligned} V(\mathbf{r}) =& \frac{1}{4\pi}\int d\mathbf{r}'\frac{\exp(-\mu|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}4\pi\frac{K}{A}\delta(\mathbf{r}') \\ =& \frac{K}{A}\frac{e^{-\mu r}}{r}, \end{aligned} \tag{4.144}$$

where $r = |\mathbf{r}|$. Substituting this into Eqn. (4.139) and rearranging for $\rho(\mathbf{r})$ gives the solution to Eqn. (4.139), which is

$$\rho(\mathbf{r}) = \frac{K}{A}\delta(\mathbf{r}) - \frac{K}{A^2}\frac{e^{-\mu r}}{r}. \tag{4.145}$$

It turns out that the above expression for $\rho(\mathbf{r})$ provides a solution to Eqn. (4.138) in the limit $aR_{\mathrm{WS}} \to \infty$. We will now show this. Consider $i \neq 0$ first. Now, the limit $aR_{\mathrm{WS}} \to \infty$ corresponds to the limit of large site volumes, or equivalently, large inter-site separations. In this limit, the $\rho(\mathbf{r})$ given by the above equation, which decays to 0 as r is increased, can be considered to be 'flat' outwith site 0 on account of the fact that the boundary of site 0, which is located at distance R_{WS} from the origin, is so far from the origin. Because of this, $\rho(\mathbf{r}) \approx \rho(\mathbf{R}_i)$ for all $\mathbf{r}$ within any site $i \neq 0$, and hence the total charge within site i is

$$\begin{aligned} Q_i =& \int_i d\mathbf{r}\ \rho(\mathbf{r}) \\ =& \rho(\mathbf{R}_i)\int_i d\mathbf{r} \\ =& \frac{4\pi}{3}R_{\mathrm{WS}}^3\rho(\mathbf{R}_i) \\ =& -\frac{4\pi R_{\mathrm{WS}}^3}{3}\frac{K}{A^2}\frac{e^{-\mu R_i}}{R_i} \end{aligned} \tag{4.146}$$

where in the last line we have used Eqn. (4.145). Similarly, the 'intra-atomic potential' at $\mathbf{R}_i$ due to the charge within site i is

$$V_i^{\mathrm{intra}} = \int_i d\mathbf{r}\frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_i|} = \rho(\mathbf{R}_i)\int_i d\mathbf{r}\frac{1}{|\mathbf{r}-\mathbf{R}_i|}. \tag{4.147}$$

This becomes

$$V_i^{\text{intra}} = 2\pi R_{\text{WS}}^2 \rho(\mathbf{R}_i) \tag{4.148}$$

after noting that

$$\begin{aligned}\int_i d\mathbf{r} \frac{1}{|\mathbf{r}-\mathbf{R}_i|} &= \int_{i=0} d\mathbf{r} \frac{1}{r} \\ &= \int_0^{R_{\text{WS}}} \frac{1}{r} 4\pi r^2 dr \\ &= 2\pi R_{\text{WS}}^2.\end{aligned} \tag{4.149}$$

Subtracting V_i^{intra} from the electrostatic potential at $\mathbf{R}_i$ gives the Madelung potential of site i:

$$\begin{aligned}V_i &= V(\mathbf{R}_i) - V_i^{\text{intra}} \\ &= -A\rho(\mathbf{R}_i) + K\delta(\mathbf{R}_i) - 2\pi R_{\text{WS}}^2 \rho(\mathbf{R}_i) \\ &= -(A + 2\pi R_{\text{WS}}^2)\rho(\mathbf{R}_i),\end{aligned} \tag{4.150}$$

where in the second line we have used Eqns. (4.139) and (4.148), and to obtain the final line we have noted that $\mathbf{R}_i \neq \mathbf{0}$ since $i \neq 0$. Rearranging Eqn. (4.146) for $\rho(\mathbf{R}_i)$ and substituting the resulting equation into the above gives

$$V_i = -aQ_i, \tag{4.151}$$

with

$$a = \frac{3A}{4\pi R_{\text{WS}}^3} + \frac{3}{2R_{\text{WS}}}. \tag{4.152}$$

Hence Eqn. (4.138) is obeyed for $i \neq 0$ in the limit $aR_{\text{WS}} \to \infty$, with a given by the above formula. Consider now the case $i = 0$. Note that $\rho(\mathbf{r})$ is *not* flat within site 0, and so we cannot proceed as above. Using Eqn. (4.145), we find that the total charge within site 0 is

$$\begin{aligned}Q_0 &= \int_{i=0} d\mathbf{r}\rho(\mathbf{r}) \\ &= \frac{K}{A}\int_{i=0} d\mathbf{r}\delta(\mathbf{r}) - \frac{K}{A^2}\int_0 d\mathbf{r}\frac{e^{-\mu r}}{r} \\ &= \frac{K}{A} - \frac{K}{A^2}\int_0^{R_{\text{WS}}} \frac{e^{-\mu r}}{r} 4\pi r^2 dr.\end{aligned} \tag{4.153}$$

Evaluating the integral, using Eqn. (4.143), and simplifying, we find that

$$Q_0 = \frac{K}{A}(1 + \mu R_{\text{WS}})e^{-\mu R_{\text{WS}}}. \tag{4.154}$$

We will now derive an expression for V_0. This can be obtained directly from Eqn. (4.145) as follows:

$$\begin{aligned} V_0 &= \sum_{i\neq 0} \int_i d\mathbf{r} \frac{\rho(\mathbf{r})}{|\mathbf{r}|} \\ &= -\frac{K}{A^2} \int_{R_{\mathrm{WS}}}^{\infty} \frac{e^{-\mu r}}{r} \frac{1}{r} 4\pi r^2 dr. \end{aligned} \tag{4.155}$$

Similarly to before, evaluating the integral, using Eqn. (4.143), and simplifying, we find that

$$V_0 = -\frac{K}{A} \mu e^{-\mu R_{\mathrm{WS}}}. \tag{4.156}$$

Now, we want to show that

$$V_0 = -aQ_0 + k_0 \tag{4.157}$$

holds. Substituting Eqns. (4.154) and (4.156) into this, it follows that Eqn. (4.157) holds if

$$k_0 = \frac{K}{A} \Big[a(1 + \mu R_{\mathrm{WS}}) - \mu \Big] e^{-\mu R_{\mathrm{WS}}}. \tag{4.158}$$

Therefore we have found a solution to Eqn. (4.138) in the limit $aR_{\mathrm{WS}} \to \infty$: Q_0 is given by Eqn. (4.154), Q_i for $i \neq 0$ is given by Eqn. (4.146), a is given by Eqn. (4.152), and k_0 is given by Eqn. (4.158).

We will now use this solution to Eqn. (4.138) to derive an expression for aG_0 which is valid in the limit $aR_{\mathrm{WS}} \to \infty$. This expression will turn out to be Eqn. (4.135). Consider a system in which $b_j = \delta_{j0}$. In this case, Eqn. (4.24) becomes

$$V_i = -aQ_i + a\delta_{i0} \tag{4.159}$$

after using Eqn. (4.25), and Eqn. (4.58) becomes

$$Q_i = aG_{i0}. \tag{4.160}$$

The above equations reveal that aG_{i0} is the charge on site i for a system in which Eqn. (4.138) holds with $k_0 = a$. If $k_0 = a$, then, from Eqn. (4.158), it follows that

$$\frac{K}{A} = a \Big[a(1 + \mu R_{\mathrm{WS}}) - \mu \Big]^{-1} e^{\mu R_{\mathrm{WS}}}. \tag{4.161}$$

Substituting this into Eqn. (4.154) gives

$$Q_0 = \frac{a(1 + \mu R_{\mathrm{WS}})}{a(1 + \mu R_{\mathrm{WS}}) - \mu} = (aG_0). \tag{4.162}$$

This can be seen to be equivalent to Eqn. (4.135) by exploiting the fact that

$$\mu = \frac{\nu}{R_{\mathrm{WS}}}, \tag{4.163}$$

which itself follows from rearranging Eqn. (4.152) for A, substituting the resulting equation into Eqn. (4.143), rearranging, and using Eqn. (4.137).

Similarly to above, we will now use the aforementioned solution to Eqn. (4.138) to derive an expression for g_β which is valid in the limit $aR_{\mathrm{WS}} \to \infty$. This expression will turn out to be Eqn. (4.136). Consider a system in which $b_j = \delta_{j0}/(aG_0)$. In this case, Eqn. (4.24) becomes

$$V_i = -aQ_i + \delta_{i0}/G_0 \tag{4.164}$$

after using Eqn. (4.25), and Eqn. (4.58) becomes

$$Q_i = G_{i0}/G_0. \tag{4.165}$$

The above equations reveal that G_{i0}/G_0 is the charge on site i for a system in which Eqn. (4.138) holds with $k_0 = 1/G_0$. From Theorem 4.1 and Eqn. (4.83), this corresponds to $Q_0 = 1$ and $Q_i = g_\beta$ for sites i at distance R_β from site 0 ($\beta > 0$). If $Q_0 = 1$, then, from Eqn. (4.154), it follows that

$$\frac{K}{A} = \frac{e^{\mu R_{\mathrm{WS}}}}{1 + \mu R_{\mathrm{WS}}}. \tag{4.166}$$

Substituting this into Eqn. (4.146) gives

$$Q_i = -\frac{4\pi R_{\mathrm{WS}}^3}{3} \frac{e^{\mu R_{\mathrm{WS}}}}{1 + \mu R_{\mathrm{WS}}} \frac{1}{A} \frac{e^{-\mu R_i}}{R_i}. \tag{4.167}$$

Using Eqn. (4.143), this becomes

$$Q_i = -\frac{1}{3} \frac{\mu^2 R_{\mathrm{WS}}^2 e^{\mu R_{\mathrm{WS}}}}{1 + \mu R_{\mathrm{WS}}} \frac{e^{-\mu R_i}}{(R_i/R_{\mathrm{WS}})}, \tag{4.168}$$

which itself becomes Eqn. (4.136) after using Eqn. (4.163) and choosing site i to be at distance R_β from site 0. ■

We can learn several things from Eqns. (4.135) and (4.136). As can be seen from Eqn. (4.137), $\nu \to 0$ as $aR_{\mathrm{WS}} \to \infty$. Applying this result to Eqn. (4.135), it can be seen that $aG_0 \to 1$ as $aR_{\mathrm{WS}} \to \infty$. Given our knowledge of the physical significance of aG_0, as was deduced in the last section, this implies that the amount of screening

vanishes as $aR_{\rm WS} \to \infty$. Consider now Eqn. (4.136). This can be written in the form

$$g_\beta = u(R_\beta/R_{\rm WS}), \tag{4.169}$$

where u is a function independent of lattice type, i.e. it is a *universal function*. Recall that g_β has the same significance as ϕ_β had in previous chapters. With this in mind, we see that the above equation describes the qualitative aspect of universal screening described in Section 2.2.2.2. To restate: universal screening is implicit in the GLCM in the limit $aR_{\rm WS} \to \infty$. We will elaborate on this result in Section 4.6.4. Another point worth mentioning regarding Eqn. (4.136) is that the function u is of Yukawa form, i.e. we can write

$$u(x) = A\frac{e^{-Bx}}{x} \tag{4.170}$$

for some constants A and B. Interestingly, the analogous parameters to g_β in the MLCM and OLCM, when parameterised using *ab initio* results in Refs. [29] and [26] respectively, were also observed to vary with $R_\beta/R_{\rm WS}$ in this manner. However, the dependence of g_β on $R_\beta/R_{\rm WS}$ observed in these studies is empirical since the values of g_β were fit to *ab initio* data. By contrast, Eqn. (4.136) is an analytical result which is increasingly accurate for higher values of $aR_{\rm WS}$, and exact in the limit $aR_{\rm WS} \to \infty$.

4.6.3 Details of Numerical Calculations

As mentioned earlier, the elements of $\mathbf{g}$ must in general be calculated numerically. We will now explain a procedure we have used to do this. The results of calculations utilising this procedure will be presented in the next section.

For a given lattice and value of a the elements of $\mathbf{g}$ can be obtained by solving the following equations:

$$(aG_0) = a\left[a + \sum_{\gamma=1}^{\infty} \frac{g_\gamma Z_\gamma}{R_\gamma}\right]^{-1}, \tag{4.171}$$

and

$$\sum_{\gamma=1}^{\infty} \Big[\Sigma_\beta(\gamma) + a\delta_{\beta\gamma}\Big] g_\gamma = -\frac{1}{R_\beta} \tag{4.172}$$

for $\beta > 0$, where $\Sigma_\beta(\gamma)$ is defined in Eqn. (3.23).

Proof: We will now derive Eqns. (4.171) and (4.172). Consider Eqn. (4.171) first. Setting $k = i$ in Eqn. (4.111) gives

$$\sum_j H_{ij} G_{ji} = 1. \tag{4.173}$$

Splitting the summation over j into contributions from each shell of i, this becomes

$$\sum_{\gamma=0}^{\infty}\sum_{j\in\gamma_i} H_{ij}G_{ji} = 1. \tag{4.174}$$

Applying Theorem 4.1 and separating out the $\gamma = 0$ term gives

$$G_0 H_{ii} + \sum_{\gamma=1}^{\infty} G_\gamma \sum_{j\in\gamma_i} H_{ij} = 1. \tag{4.175}$$

Noting that $H_{ii} = a$ and $H_{ij} = 1/R_\gamma$ if $j \in \gamma_i$ and $\gamma \geq 1$ (which follows from Eqns. (4.28) and (2.55)), the left-hand side of the above equation becomes

$$G_0 a + \sum_{\gamma=1}^{\infty} G_\gamma \sum_{j\in\gamma_i} \frac{1}{R_\gamma} = G_0 a + \sum_{\gamma=1}^{\infty} \frac{G_\gamma}{R_\gamma} \sum_{j\in\gamma_i} 1 = G_0 a + \sum_{\gamma=1}^{\infty} \frac{G_\gamma Z_\gamma}{R_\gamma}, \tag{4.176}$$

where we have used Eqn. (4.89). Therefore

$$G_0 a + \sum_{\gamma=1}^{\infty} \frac{G_\gamma Z_\gamma}{R_\gamma} = 1, \tag{4.177}$$

which becomes

$$G_0 \left[a + \sum_{\gamma=1}^{\infty} \frac{g_\gamma Z_\gamma}{R_\gamma} \right] = 1 \tag{4.178}$$

after using Eqn. (4.83) and factorising. Rearranging this for G_0 and multiplying the resulting equation by a gives Eqn. (4.171).

We will now derive Eqn. (4.172). Consider Eqn. (4.117). Note that this holds only for $\beta > 0$. Using Eqn. (3.22), this can be seen to be equivalent to

$$\sum_{\gamma=1}^{\infty} \left[\frac{Z_\gamma}{R_\beta} - \Sigma_\beta(\gamma) \right] G_\gamma = a G_\beta. \tag{4.179}$$

Dividing both sides by G_0 and using Eqn. (4.83) this becomes

$$\sum_{\gamma=1}^{\infty} \left[\frac{Z_\gamma}{R_\beta} - \Sigma_\beta(\gamma) \right] g_\gamma = a g_\beta, \tag{4.180}$$

which in turn becomes

$$\frac{1}{R_\beta} \sum_{\gamma=1}^{\infty} Z_\gamma g_\gamma - \sum_{\gamma=1}^{\infty} \Sigma_\beta(\gamma) g_\gamma = a g_\beta \tag{4.181}$$

after expanding the summation. Substituting Eqn. (4.84) into the first term on the left-hand side, and applying the sifting property of the Kronecker delta to the right-hand side, we find that

$$-\frac{1}{R_\beta}-\sum_{\gamma=1}^{\infty}\Sigma_\beta(\gamma)g_\gamma=a\sum_{\gamma=1}^{\infty}\delta_{\beta\gamma}g_\gamma, \tag{4.182}$$

which becomes Eqn. (4.172) after moving the second term on the left-hand side to the right-hand side and factorising. ■

Eqns. (4.171) and (4.172) can be solved numerically using the following procedure:

1. Determine the values of $1/R_\beta$ and $\Sigma_\beta(\gamma)$ for the lattice in question.
2. Solve the infinite system of linear equations in g_β described by Eqn. (4.172) using the relevant value of a.
3. Substitute the values of g_β obtained from the last step into Eqn. (4.171) to obtain (aG_0).

The second step is problematic since an infinite system of linear equations cannot be solved exactly using numerical methods. We can overcome this problem by truncating the summation in Eqn. (4.172) at $\gamma = t$ and considering only $\beta \leq t$. In this case we have a system of t linear equations to solve:

$$\sum_{\gamma=1}^{t}\Big[\Sigma_\beta(\gamma)+a\delta_{\beta\gamma}\Big]g_\gamma=-\frac{1}{R_\beta}\quad\text{for } 1\leq\beta\leq t. \tag{4.183}$$

Their solution can be found by using conventional linear algebra algorithms (e.g. Gaussian elimination), giving us the values of g_β for $1 \leq \beta \leq t$. Is the aforementioned truncation justified? *Under the assumption that* $g_\beta < 0$ *for all* $\beta > t$, a sufficient condition for the truncation to be justified is if

$$\frac{1}{R_1}\Big(1+\sum_{\gamma=1}^{t}Z_\gamma g_\gamma\Big)\leq\epsilon, \tag{4.184}$$

where ϵ is a number small enough to be considered 0 for practical purposes.

Proof: By comparing Eqn. (4.183) to Eqn. (4.172) it can be seen that the former is valid with regards to determining the values of g_β for $1 \leq \beta \leq t$ if

$$\left|\sum_{\gamma=t+1}^{\infty}\Sigma_\beta(\gamma)g_\gamma\right|\leq\epsilon, \tag{4.185}$$

i.e. the sum of the terms not included in the left-hand side of Eqn. (4.183) is negligible. Substituting Eqn. (3.23) into the above gives

$$\left| \sum_{\gamma=t+1}^{\infty} \sum_{\delta=1}^{\infty} \frac{K_\beta^\delta(\gamma)}{R_\delta} g_\gamma \right| \leq \epsilon. \tag{4.186}$$

Noting that the values of $K_\beta^\delta(\gamma)$ and R_δ are all non-negative, it must be the case that the above inequality is true if the inequality

$$\sum_{\gamma=t+1}^{\infty} \sum_{\delta=1}^{\infty} \frac{K_\beta^\delta(\gamma)}{R_\delta} |g_\gamma| \leq \epsilon \tag{4.187}$$

is true. Noting that $1/R_\delta \leq 1/R_1$ for all $\delta > 0$, it can be seen that the above inequality is in turn true if

$$\sum_{\gamma=t+1}^{\infty} \frac{1}{R_1} |g_\gamma| \sum_{\delta=1}^{\infty} K_\beta^\delta(\gamma) \leq \epsilon \tag{4.188}$$

is true. Now, from the definition of $K_\beta^\delta(\gamma)$, the following identity holds (for all β):

$$\sum_{\delta=0}^{\infty} K_\beta^\delta(\gamma) = Z_\gamma. \tag{4.189}$$

Moving the $\delta = 0$ term to the right-hand side, and substituting the resulting equation into the above inequality, gives

$$\frac{1}{R_1} \sum_{\gamma=t+1}^{\infty} |g_\gamma| (Z_\gamma - K_\beta^0(\gamma)) \leq \epsilon. \tag{4.190}$$

This becomes

$$\frac{1}{R_1} \sum_{\gamma=t+1}^{\infty} |g_\gamma| Z_\gamma \leq \epsilon \tag{4.191}$$

after noting that $\gamma > t$ and $\beta \leq t$, and hence - from the definition of $K_\beta^\delta(\gamma)$ - $K_\beta^0(\gamma) = 0$. Applying our assumption that $g_\gamma < 0$ for all $\gamma > t$, it follows that

$$-\frac{1}{R_1} \sum_{\gamma=t+1}^{\infty} g_\gamma Z_\gamma \leq \epsilon. \tag{4.192}$$

Finally, substituting the equation

$$\sum_{\gamma=t+1}^{\infty} g_\gamma Z_\gamma = -1 - \sum_{\gamma=1}^{t} Z_\gamma g_\gamma \tag{4.193}$$

- which follows from Eqn. (4.84) - into the above inequality gives Eqn. (4.184). ■

With the above in mind, a practical implementation of step 2 is as follows:

1. Decide on values of t and ϵ.
2. Solve the set of t equations described by Eqn. (4.183) to obtain g_β for $1 \leq \beta \leq t$.
3. Check that the values of g_β obtained from the last step are such that Eqn. (4.184) is true.
4. Check that the assumption that $g_\beta < 0$ for $\beta > t$ (under which Eqn. (4.184) is derived) looks reasonable. We will assume that if $|g_\beta|$ is monotonically decreasing and $g_\beta < 0$ 'for many β' before $\beta = t$, then it will continue to do so for $\beta > t$, in which case the said assumption is reasonable.

If the values of g_β fail any of the 'checks' described above then one cannot assume that the calculated values of g_β are accurate. In this case one may wish to recalculate the values of g_β with increasing values of t until the checks are passed. This is not done in our calculations: we simply do not present results which fail the checks.

We have used the above procedure to obtain $\mathbf{g}$ for a at 0.01 intervals between 0 and 7 for a fcc, bcc and sc lattice, each chosen to have $R_{\mathrm{WS}} = 1$. In all calculations we used $t = 100$ and $\epsilon = 1 \times 10^{-7}$. The system of equations described by Eqn. (4.183) was solved numerically using subroutines from the LAPACK library [67]. Specifically, the equations were solved using a LU decomposition with partial pivoting using LAPACK's *DGETRF* and *DGETRS* subroutines. The results of these calculations essentially gives $\mathbf{g}$ as a function of a, and hence, since $R_{\mathrm{WS}} = 1$, aR_{WS} for each of the said lattices.

4.6.4 Results of Numerical Calculations and Discussion

The results of the aforementioned numerical calculations are presented graphically in Figs. 4.1 and 4.2, where the analytical results of Eqns. (4.135) and (4.136) are also shown for comparison. Fig. 4.1 shows how aG_0 varies with aR_{WS}, and Fig. 4.2 shows the values of g_β plotted against R_β/R_{WS} at selected values of aR_{WS}. Note that in both figures the numerical results tend to the analytical results (the blue curves) as aR_{WS} is increased. This is expected because the analytical results are increasingly accurate for higher values of aR_{WS}. We will begin by discussing Fig. 4.1. This figure reveals that, for all lattice types considered, aG_0 decreases monotonically to 1 - the limit predicted earlier by Eqn. (4.135) - as aR_{WS} increases. Recall that aG_0 is a measure of the amount of screening. Therefore the amount of screening decreases monotonically with aR_{WS}. This makes sense. Earlier, we mentioned that the quantity aR_{WS} determines

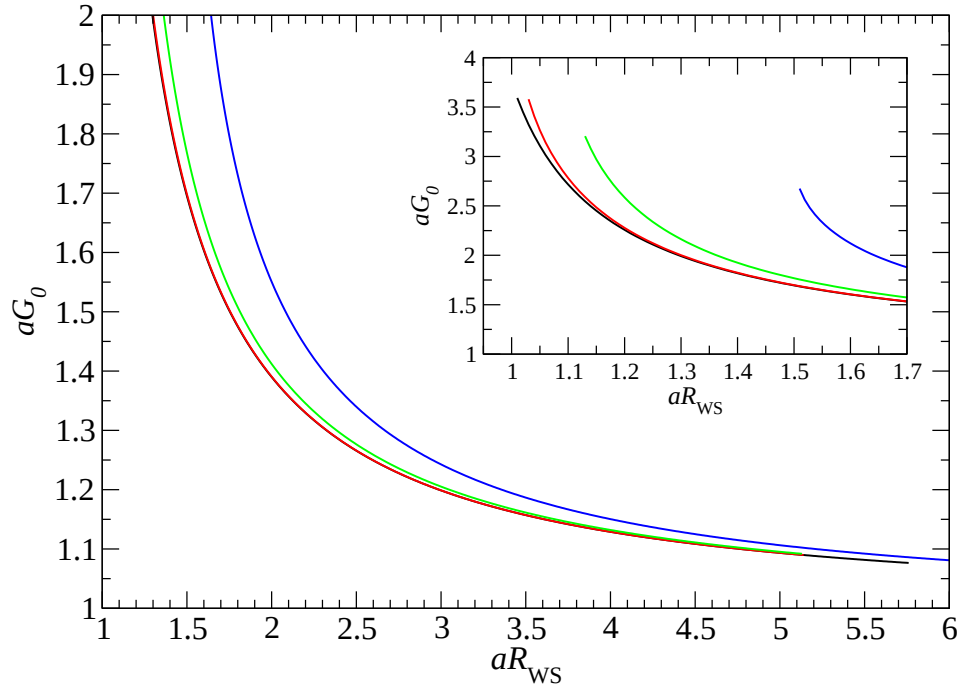


Figure 4.1: aG_0 vs. aR_{WS}. The black, red and green curves correspond to the results of the numerical calculations for the fcc, bcc and sc lattices respectively. The blue curve corresponds to the predictions of Eqn. (4.135).

the strength of the local interactions relative to the strength of the inter-site (Coulomb) interactions. Bearing in mind that screening arises wholly as an attempt to reduce the energy associated with inter-site interactions, it follows that for high values of aR_{WS}, i.e. weak inter-site interactions, the amount of screening will be low. Conversely, for low values of aR_{WS}, the amount of screening will be high, which is also borne out in the figure: aG_0 diverges as aR_{WS} is decreased to 0.

Consider now Fig. 4.2. Recall that g_β determines the amount of screening charge which is located on a site in shell β of a charge perturbation. The tendency for the screening to decrease as aR_{WS} increases is reflected in this figure: at high values of aR_{WS} the radial distribution of screening charge[11] is 'flat', which corresponds to less screening; while for lower values of aR_{WS} the radial distribution is increasingly 'skewed' towards

[11]Note that the term 'radial distribution' here refers to the distribution of screening charge outwith the site hosting a charge perturbation. Thus the term has a different significance than in Section 4.1, where it referred to the distribution of electrons within a site.

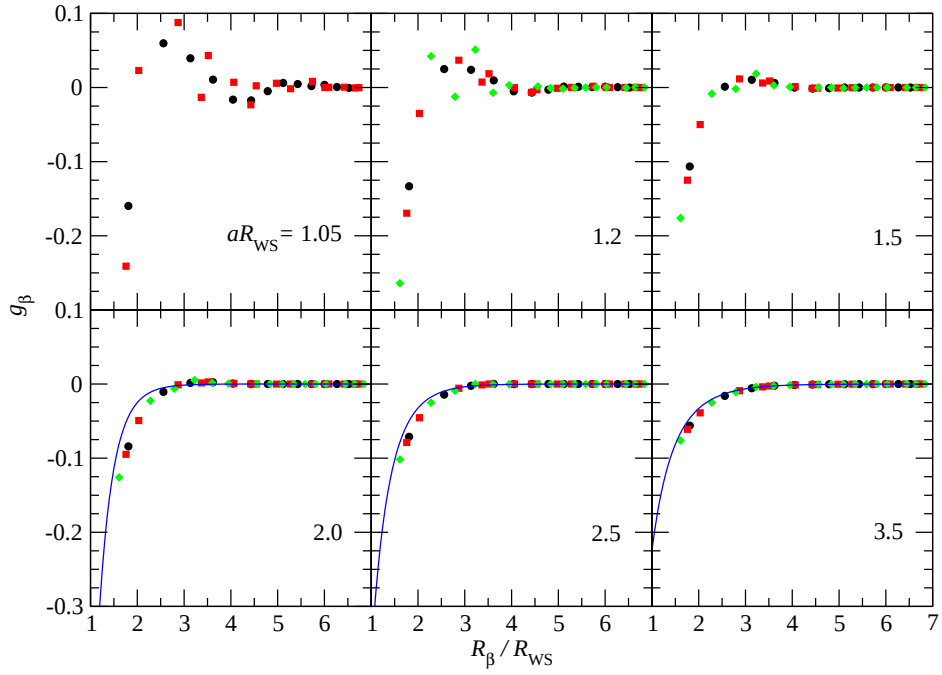


Figure 4.2: g_β vs. $R_\beta/R_{\rm WS}$ for various values of $aR_{\rm WS}$. Each panel corresponds to a particular $aR_{\rm WS}$, whose value is shown in the label. The black circles, red squares and green diamonds correspond to the results of the numerical calculations for the fcc, bcc and sc lattices respectively. The blue curves correspond to the predictions of Eqn. (4.136).

the central site, which corresponds to more screening.[12] Interestingly, while the values of g_β increase to 0 monotonically for large values of aR_{WS}, they exhibit oscillations for lower values of aR_{WS}. Furthermore, the oscillations become increasingly violent as aR_{WS} is decreased. This is a Gibb's-phenomenon-like effect. Recall that Eqn. (4.136) describes how g_β depends on R_β/R_{WS} for large values of aR_{WS}, and is of Yukawa form. The range of the g_β vs. R_β/R_{WS} curve of Eqn. (4.136) is determined by the value of aR_{WS}: the higher aR_{WS} is, the longer its range in real space, and hence the shorter its range in reciprocal space. Now, due to the discreteness of the lattice, there is an upper limit to the wavevectors which can be represented upon it. This is important for low values of aR_{WS}. Here, the Yukawa function has strong Fourier components above the lattice's upper limit, the result of which is that these components are 'cut out' of the Yukawa function, leaving oscillations in the g_β vs. R_β/R_{WS} curve in real space. Note that the upper limit is different for different lattices, which is why the g_β vs. R_β/R_{WS} curves at low values of aR_{WS} differ for different lattice types. Now, as aR_{WS} is increased, the Fourier components above the lattice's upper limit become weaker, and hence the amount of components cut out of the Yukawa function decreases. Because of this, the g_β vs. R_β/R_{WS} curves for all lattices increasingly take the form of the Yukawa function, i.e. the g_β vs. R_β/R_{WS} curve increasingly becomes universal, as aR_{WS} is increased. Conversely, the universality increasingly breaks down as aR_{WS} is decreased from infinity. This is borne out in Fig. 4.2. Consider the panel in the figure corresponding to $aR_{\mathrm{WS}} = 1.5$. Here we see that, while the $(R_\beta/R_{\mathrm{WS}}, g_\beta)$ points for the fcc and bcc lattices still appear to lie upon the same curve, the points for the sc lattice do not. Decreasing aR_{WS} below this only increases the deviation of the sc points from the fcc/bcc curve. At $aR_{\mathrm{WS}} = 1.2$ the deviation is significant. Further decreases in aR_{WS} similarly cause the bcc curve to 'break away' from the fcc curve, as can be seen from the $aR_{\mathrm{WS}} = 1.05$ panel in the figure.

We will now compare our results to those of RS described in Section 2.2.2.2. Recall that RS, using the SSLSGF method, observed that for all systems g_β[13] vs. R_β/R_{WS} is a universal curve. They also observed that $a_X R_{\mathrm{WS}} \approx 1.6$ for all species in all systems. We expect our results to be in quantitative agreement with those of RS, because the approximations which underpin the GLCM are either implicit in the SSLSGF method or can be justified *a posteriori* from RS's results themselves. To elaborate, there are four approximations which underpin the GLCM. The first three of these are listed in

[12] To clarify, by 'less screening' and 'more screening' we are referring to the *distribution* of screening charge associated with a charge perturbation - whose charge we will denote here as δQ - and not the *amount* of screening charge, which is always $-\delta Q$. 'Less screening' means that the $-\delta Q$ is distributed further from the perturbation; 'more screening' means that it is distributed closer to the perturbation.

[13] In Section 2.2.2.2, the quantity ϕ_β is referred to instead of g_β. This is of no consequence because, as was discussed in Section 4.6.1, g_β has the same physical significance as ϕ_β.

Section 4.1.2. Firstly, there is the spherical approximation. This is utilised in the SSLSGF method. Secondly, there is the 'local approximation'. This is also utilised in the SSLSGF method through the effective medium approach described in Section 4.1.2: as mentioned in Section 2.2.2.2, in the SSLSGF method, each site 'sees' the LSGF effective medium as its surroundings. Thirdly, there is the approximation that the values of $Q_i - b_i$ for all i are small. It was pointed out in Section 4.1.2 that if the aforementioned three approximations are satisfied, then the Q-V relations *must* hold. The fact that the Q-V relations are observed to hold to a high degree of accuracy in RS's results [24] therefore implies that the values of $Q_i - b_i$ for all i must be small. RS's results also give justification for the final approximation which underpins the GLCM, namely, that the values of a_X are the same for all species in a given system: RS's results reveal that $a_X R_{\mathrm{WS}} \approx 1.6$, for which it is necessary that the values of a_X are the same for all species within any particular system. Note that the exact reasons *why* the values of a_X are the same for all species in a given system, and are such that $a_X R_{\mathrm{WS}} \approx 1.6$ for all systems, are still unclear and require further investigation. In Fig. 4.3, our results for $aR_{\mathrm{WS}} = 1.6$ are compared to the curve obtained by RS. From the figure, it can be seen that our fcc and bcc points agree well with the curve - as expected. However, the points corresponding to the sc lattice do not. This is because RS only considered systems with the fcc, bcc or bct structure: they did not consider systems with the sc structure. An alloy exhibiting the sc structure would therefore 'break' the universality described by RS.

4.7 Analytical Expressions for Various Physical Quantities

We will now use Eqn. (4.96) to derive analytical expressions for various physical quantities. To be as general as possible, the expressions which we will derive will pertain to any set of sites S within the system under consideration, e.g. S could be the set of all sites in the entire system, all sites in a particular localised region, or all sites in a particular layer.

We begin by introducing the S-dependent quantities p_{β}^{XY} and $p_{\beta\gamma}^{XYZ}$, which will appear in the expressions derived below. These are defined by the equations

$$p_{\beta}^{XY} = \langle N_{jY0} \rangle_{j \in \beta_i \,, i \in X \cap S} \tag{4.194}$$

and

$$p_{\beta\gamma}^{XYZ} = \langle N_{jY0} N_{kZ0} \rangle_{j \in \beta_i \,, k \in \gamma_i \setminus j \,, i \in X \cap S}. \tag{4.195}$$

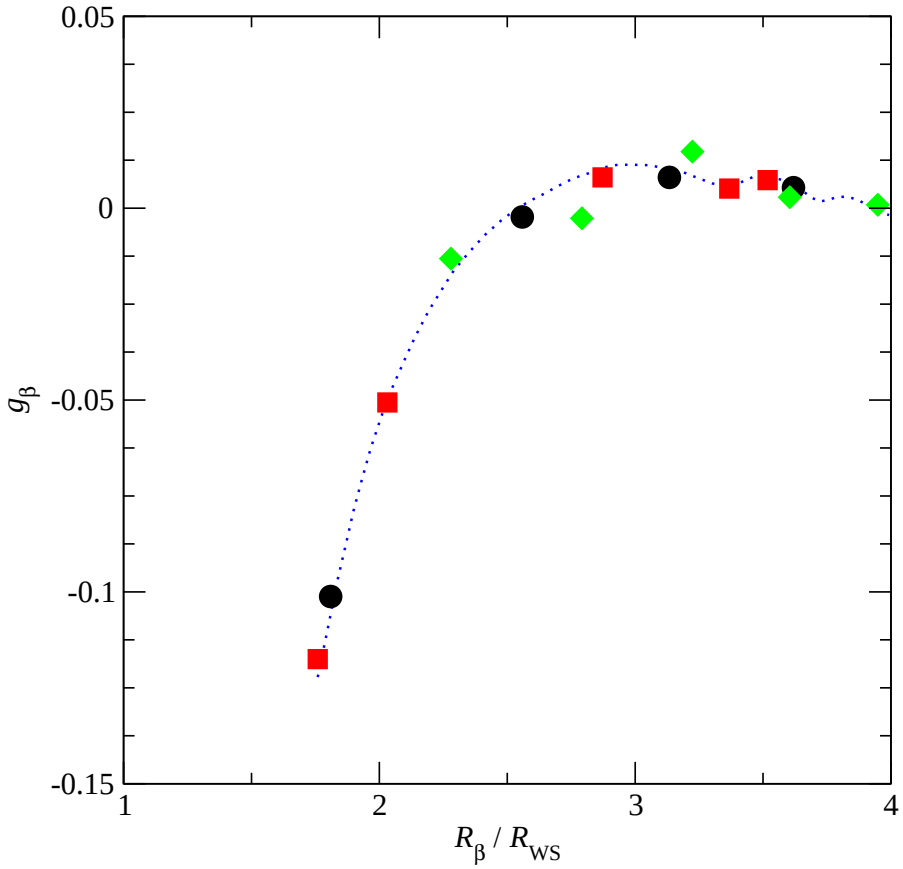


Figure 4.3: The values of g_β plotted against R_β/R_{WS} for the fcc (black circles), bcc (red squares) and sc (green diamonds) for $aR_{\mathrm{WS}} = 1.6$. The dotted blue curve corresponds to the results obtained by RS in Ref. [24].

Note that, as follows from the fact that shell 0 of site j consists of only site j, the quantity N_{jY0} is 1 if site j belongs to species Y and 0 otherwise. Note also that, while the parameters p_β^{XY} and $p_{\beta\gamma}^{XYZ}$ depend on S, we have not indicated this explicitly for the sake of notational clarity. The physical significance of the quantity p_β^{XY} is as follows: p_β^{XY} is the probability that a site at distance R_β from an X site in S belongs to species Y - where both sites are selected at random within the aforementioned constraints. It should be noted the values of p_β^{XY} for a given X and β are constrained to obey the equation

$$\sum_Y p_\beta^{XY} = 1, \tag{4.196}$$

i.e. the probability that a site at distance R_β from a randomly selected X site within S belongs to *any* species is 1. Furthermore, if the system under consideration is a binary alloy, and S is the set of *all* sites in the system, then the values of p_β^{XY} are related to α_β - the Warren-Cowley short range order parameter pertaining to sites separated by R_β - by the following equations [68]:

$$p_\beta^{AA} = c_A + c_B\alpha_\beta, \tag{4.197}$$

$$p_\beta^{AB} = c_B(1 - \alpha_\beta), \tag{4.198}$$

$$p_\beta^{BA} = c_A(1 - \alpha_\beta), \tag{4.199}$$

$$p_\beta^{BB} = c_B + c_A\alpha_\beta. \tag{4.200}$$

In a similar vein to p_β^{XY}, $p_{\beta\gamma}^{XYZ}$ is the probability that sites j and k, which are distinct and at distances R_β and R_γ respectively from an X site in S, belong to species Y and Z respectively - where all sites are selected randomly within the aforementioned constraints. Note that by saying *distinct* we have disallowed the case where j and k are the same site (which could only occur if $\gamma = \beta$). Similarly to p_β^{XY}, the values of $p_{\beta\gamma}^{XYZ}$ for a given X, β and γ are constrained to obey the equation

$$\sum_{Y,Z} p_{\beta\gamma}^{XYZ} = 1. \tag{4.201}$$

4.7.1 Charges

The first expressions which we will derive are for the mean and variance in Q_i for X sites in S, which we denote as $\langle Q\rangle_X^S$ and $\mathrm{Var}(Q)_X^S$ respectively, i.e. $\langle Q\rangle_X^S = \langle Q_i\rangle_{i\in X\cap S}$ and $\mathrm{Var}(Q)_X^S = \mathrm{Var}(Q_i)_{i\in X\cap S}$. A similar notation will be used for the mean and variance

in V_i and $\Delta E_i^{\mathrm{B,i}}$ later. These expressions are

$$\langle Q\rangle_X^S = aG_0 \sum_Y b_{YX} \sum_{\beta=1}^{\infty} g_\beta Z_\beta p_\beta^{XY} \tag{4.202}$$

and

$$\begin{aligned}\mathrm{Var}(Q)_X^S =&(aG_0)^2 \Bigg\{ \sum_Y b_{YX}^2 \sum_{\beta=1}^{\infty} g_\beta^2 Z_\beta p_\beta^{XY} - \sum_{Y,Z} b_{YX} b_{ZX} \sum_{\beta=1}^{\infty} g_\beta^2 Z_\beta p_{\beta\beta}^{XYZ} \\ &+ \sum_{Y,Z} b_{YX} b_{ZX} \sum_{\beta=1}^{\infty}\sum_{\gamma=1}^{\infty} g_\beta g_\gamma Z_\beta Z_\gamma (p_{\beta\gamma}^{XYZ} - p_\beta^{XY} p_\gamma^{XZ}) \Bigg\}.\end{aligned} \tag{4.203}$$

Proof: We will now derive Eqns. (4.202) and (4.203), beginning with Eqn. (4.202). Consider Eqn. (4.96). Taking the mean of this over X sites in S gives

$$\langle Q\rangle_X^S = aG_0 \sum_Y b_{YX} \sum_{\beta=1}^{\infty} g_\beta \langle N_{iY\beta}\rangle_{i\in X\cap S}, \tag{4.204}$$

where we have exploited the linearity of the mean (Eqn. (3.17)). Now, by definition,

$$\langle N_{iY\beta}\rangle_{i\in X\cap S} = \frac{1}{|X\cap S|} \sum_{i\in X\cap S} N_{iY\beta}. \tag{4.205}$$

Using the fact that

$$N_{iY\beta} = \sum_{j\in\beta_i} N_{jY0}, \tag{4.206}$$

the above expression for $\langle N_{iY\beta}\rangle_{i\in X\cap S}$ becomes

$$\langle N_{iY\beta}\rangle_{i\in X\cap S} = \frac{1}{|X\cap S|} \sum_{i\in X\cap S} \sum_{j\in\beta_i} N_{jY0}, \tag{4.207}$$

which in turn becomes

$$\begin{aligned}\langle N_{iY\beta}\rangle_{i\in X\cap S} =& Z_\beta \left(\frac{1}{|X\cap S| Z_\beta} \sum_{i\in X\cap S} \sum_{j\in\beta_i} N_{jY0} \right) \\ =& Z_\beta \langle N_{jY0}\rangle_{j\in\beta_i, i\in X\cap S},\end{aligned} \tag{4.208}$$

where we have used the definition of $\langle N_{jY0}\rangle_{j\in\beta_i,i\in X\cap S}$. Comparing the above equation to Eqn. (4.194), it follows that

$$\langle N_{iY\beta}\rangle_{i\in X\cap S} = Z_\beta p_\beta^{XY}, \tag{4.209}$$

which, when substituted into Eqn. (4.204), gives Eqn. (4.202).

We will now derive Eqn. (4.203). To do this, we require the following property of the variance: for some set of properties $P_i^1, P_i^2, \ldots, P_i^{m_{\max}}$ pertaining to each site i in a set of sites T whose variance we are interested in, the variance of the quantity $B + \sum_{m=1}^{m_{\max}} A^m P_i^m$, where $A^1, A^2, \ldots, A^{m_{\max}}$ and B are constants, obeys

$$\mathrm{Var}\left(B + \sum_{m=1}^{m_{\max}} A^m P_i^m\right)_{i\in T} = \sum_{m=1}^{m_{\max}} \sum_{n=1}^{m_{\max}} A^m A^n \,\mathrm{Cov}(P_i^m, P_i^n)_{i\in T}, \tag{4.210}$$

where

$$\mathrm{Cov}(P_i^m, P_i^n)_{i\in T} = \langle P_i^m P_i^n\rangle_{i\in T} - \langle P_i^m\rangle_{i\in T}\langle P_i^n\rangle_{i\in T} \tag{4.211}$$

is the covariance between the properties P^m and P^n for sites in T. Taking the variance of Eqn. (4.96) for X sites in S, and applying Eqn. (4.210), we find that

$$\mathrm{Var}(Q)_X^S = \sum_Y \sum_{\beta=1}^{\infty} \sum_Z \sum_{\gamma=1}^{\infty} (aG_0 b_{YX} g_\beta)(aG_0 b_{ZX} g_\gamma)\,\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S}. \tag{4.212}$$

This becomes

$$\mathrm{Var}(Q)_X^S = (aG_0)^2 \sum_{Y,Z} b_{YX} b_{ZX} \sum_{\beta=1}^{\infty} \sum_{\gamma=1}^{\infty} g_\beta g_\gamma \,\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S} \tag{4.213}$$

after changing the order of the summations.

We will now derive an expression for $\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S}$ which, when substituted into the above equation, gives Eqn. (4.203). From the definition of the covariance (Eqn. (4.211)),

$$\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S} = \langle N_{iY\beta} N_{iZ\gamma}\rangle_{i\in X\cap S} - \langle N_{iY\beta}\rangle_{i\in X\cap S}\langle N_{iZ\gamma}\rangle_{i\in X\cap S}. \tag{4.214}$$

Substituting Eqn. (4.209) into this gives

$$\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S} = \langle N_{iY\beta} N_{iZ\gamma}\rangle_{i\in X\cap S} - Z_\beta Z_\gamma p_\beta^{XY} p_\gamma^{XZ}. \tag{4.215}$$

Consider now $\langle N_{iY\beta} N_{iZ\gamma}\rangle_{i\in X\cap S}$. By definition, this is

$$\langle N_{iY\beta} N_{iZ\gamma}\rangle_{i\in X\cap S} = \frac{1}{|X\cap S|} \sum_{i\in X\cap S} N_{iY\beta} N_{iZ\gamma}, \tag{4.216}$$

which becomes

$$\langle N_{iY\beta}N_{iZ\gamma}\rangle_{i\in X\cap S} = \frac{1}{|X\cap S|}\sum_{i\in X\cap S}\sum_{j\in\beta_i}\sum_{k\in\gamma_i} N_{jY0}N_{kZ0} \tag{4.217}$$

after using Eqn. (4.206). For $\gamma \neq \beta$, the above equation becomes

$$\begin{aligned}\langle N_{iY\beta}N_{iZ\gamma}\rangle_{i\in X\cap S} &= Z_\beta Z_\gamma\left(\frac{1}{|X\cap S|Z_\beta Z_\gamma}\sum_{i\in X\cap S}\sum_{j\in\beta_i}\sum_{k\in\gamma_i} N_{jY0}N_{kZ0}\right)\\ &= Z_\beta Z_\gamma\langle N_{jY0}N_{kZ0}\rangle_{j\in\beta_i\,,\,k\in\gamma_i\setminus j\,,\,i\in X\cap S},\end{aligned} \tag{4.218}$$

where we have used the definition of $\langle N_{jY0}N_{kZ0}\rangle_{j\in\beta_i\,,\,k\in\gamma_i,\,i\in X\cap S}$, as well as the fact that: if j is a site in shell β of i, k is a site in shell γ of i, and $\gamma\neq\beta$, then it cannot be the case that $k=j$, i.e. the sets γ_i and $\gamma_i\setminus j$ are equivalent. Comparing Eqn. (4.218) to Eqn. (4.195) gives

$$\langle N_{iY\beta}N_{iZ\gamma}\rangle_{i\in X\cap S} = Z_\beta Z_\gamma p_{\beta\gamma}^{XYZ}, \tag{4.219}$$

which we will use in a moment. For $\gamma=\beta$, Eqn. (4.217) becomes

$$\langle N_{iY\beta}N_{iZ\beta}\rangle_{i\in X} = \frac{1}{|X\cap S|}\sum_{i\in X\cap S}\sum_{j\in\beta_i}\left[N_{jY0}N_{jZ0} + \sum_{k\in\beta_i\setminus j} N_{jY0}N_{kZ0}\right] \tag{4.220}$$

after separating out the $k=j$ term from the summation over k. Expanding the above equation, and noting that $N_{jY0}N_{jZ0} = \delta_{YZ}N_{jY0}$ (where δ_{YZ} denotes the Kronecker delta: $\delta_{YZ}=1$ if $Y=Z$ and 0 otherwise), gives

$$\begin{aligned}\langle N_{iY\beta}N_{iZ\beta}\rangle_{i\in X\cap S} =& \delta_{YZ}\frac{1}{|X\cap S|}\sum_{i\in X\cap S}\sum_{j\in\beta_i} N_{jY0} + \frac{1}{|X\cap S|}\sum_{i\in X\cap S}\sum_{j\in\beta_i}\sum_{k\in\beta_i\setminus j} N_{jY0}N_{kZ0}\\ =& Z_\beta\delta_{YZ}\left(\frac{1}{|X\cap S|Z_\beta}\sum_{i\in X\cap S}\sum_{j\in\beta_i} N_{jY0}\right)\\ &+ Z_\beta(Z_\beta-1)\left(\frac{1}{|X\cap S|Z_\beta(Z_\beta-1)}\sum_{i\in X\cap S}\sum_{j\in\beta_i}\sum_{k\in\beta_i\setminus j} N_{jY0}N_{kZ0}\right).\end{aligned} \tag{4.221}$$

Noting that, by definition, the quantities in parenthesis in the first and second terms are $\langle N_{jY0}\rangle_{,j\in\beta_i\,,\,i\in X\cap S}$ and $\langle N_{jY0}N_{kZ0}\rangle_{j\in\beta_i\,,\,k\in\beta_i\setminus j\,,\,i\in X\cap S}$ respectively, we obtain

$$\begin{aligned}\langle N_{iY\beta}N_{iZ\beta}\rangle_{i\in X\cap S} =& Z_\beta\delta_{YZ}\langle N_{jY0}\rangle_{,j\in\beta_i\,,\,i\in X\cap S}\\ &+ Z_\beta(Z_\beta-1)\langle N_{jY0}N_{kZ0}\rangle_{j\in\beta_i\,,\,k\in\beta_i\setminus j\,,\,i\in X\cap S},\end{aligned} \tag{4.222}$$

which in turn becomes

$$\begin{aligned}\langle N_{iY\beta}N_{iZ\beta}\rangle_{i\in X\cap S} =& Z_\beta\delta_{YZ}p_\beta^{XY} + Z_\beta(Z_\beta-1)p_{\beta\beta}^{XYZ}\\ =& Z_\beta Z_\beta p_{\beta\beta}^{XYZ} + Z_\beta\delta_{YZ}p_\beta^{XY} - Z_\beta p_{\beta\beta}^{XYZ}.\end{aligned} \tag{4.223}$$

after using Eqns. (4.194) and (4.195). Comparing the above to Eqn. (4.219) - which holds for $\gamma \neq \beta$ - it can be seen that

$$\langle N_{iY\beta}N_{iZ\gamma}\rangle_{i\in X\cap S} = Z_\beta Z_\gamma p_{\beta\gamma}^{XYZ} + \delta_{\beta\gamma}Z_\beta(\delta_{YZ}p_\beta^{XY} - p_{\beta\beta}^{XYZ}) \tag{4.224}$$

holds for all γ. Substituting this into Eqn. (4.215) gives the expression for $\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S}$ which we have been seeking:

$$\mathrm{Cov}(N_{iY\beta}, N_{iZ\gamma})_{i\in X\cap S} = Z_\beta Z_\gamma(p_{\beta\gamma}^{XYZ} - p_\beta^{XY}p_\gamma^{XZ}) + \delta_{\beta\gamma}Z_\beta(\delta_{YZ}p_\beta^{XY} - p_{\beta\beta}^{XYZ}). \tag{4.225}$$

Now, the above equation, when substituted into Eqn. (4.213), gives

$$\begin{aligned}\mathrm{Var}(Q)_X^S =& (aG_0)^2\Bigg\{\sum_{Y,Z}b_{YX}b_{ZX}\sum_{\beta=1}^{\infty}\sum_{\gamma=1}^{\infty}g_\beta g_\gamma Z_\beta Z_\gamma(p_{\beta\gamma}^{XYZ} - p_\beta^{XY}p_\gamma^{XZ})\\ &+\sum_{Y,Z}b_{YX}b_{ZX}\sum_{\beta=1}^{\infty}\sum_{\gamma=1}^{\infty}g_\beta g_\gamma\delta_{\beta\gamma}Z_\beta(\delta_{YZ}p_\beta^{XY} - p_{\beta\beta}^{XYZ})\Bigg\}.\end{aligned} \tag{4.226}$$

Eqn. (4.203) is retrieved from the above by expanding the second term, using the sifting property of the Kronecker delta, and then rearranging the resulting equation. ■

4.7.2 Madelung Potentials

If $\langle Q\rangle_X^S$ and $\mathrm{Var}(Q)_X^S$ are known, $\langle V\rangle_X^S$ and $\mathrm{Var}(V)_X^S$ can be calculated using the following equations:

$$\langle V\rangle_X^S = -a\langle Q\rangle_X^S + ab_X \tag{4.227}$$

and

$$\mathrm{Var}(V)_X^S = a^2\,\mathrm{Var}(Q)_X^S. \tag{4.228}$$

Proof: We will now derive Eqns. (4.227) and (4.228). Consider Eqn. (4.24), which applies to all X sites. Taking the mean of this over all X sites in S, and exploiting the linearity of the mean (Eqn. (3.17)) gives

$$\langle V_i\rangle_{i\in X\cap S} = -a\langle Q_i\rangle_{i\in X\cap S} + k_X. \tag{4.229}$$

This becomes Eqn. (4.227) after using Eqn. (4.26). Taking the variance of Eqn. (4.24) over all X sites in S gives Eqn. (4.228) after using Eqn. (4.210). ■

4.7.3 Initial State Core Level Shifts

Using Eqn. (2.88), we can derive similar equations to Eqns. (4.227) and (4.228) for the CLSs.[14] Assuming that r_i^{eff} takes the same value r_X^{eff} for all X sites, these expressions are

$$\langle \Delta E^{\text{B,i}} \rangle_X^S = \left(\frac{1}{r_X^{\text{eff}}} - a \right) \langle Q \rangle_X^S + a b_X + \Theta_X \tag{4.230}$$

and

$$\text{Var}(\Delta E^{\text{B,i}})_X^S = \left(\frac{1}{r_X^{\text{eff}}} - a \right)^2 \text{Var}(Q)_X^S, \tag{4.231}$$

where

$$\Theta_X = Q_{\text{ref},X}^{\text{val}} \left(\frac{1}{r_X^{\text{eff}}} - \frac{1}{r_{\text{ref},X}^{\text{eff}}} \right), \tag{4.232}$$

$Q_{\text{ref},X}^{\text{val}}$ is the amount of valence charge on any site in a pure X metal, and $r_{\text{ref},X}^{\text{eff}}$ is the effective radius associated with a pure X metal.

Proof: We will now derive Eqns. (4.230) and (4.231). Note that, if $r_i^{\text{eff}} = r_X^{\text{eff}}$, then the final term in Eqn. (2.88) becomes Θ_X as defined above. With this in mind, substituting $r_i^{\text{eff}} = r_X^{\text{eff}}$ into Eqn. (2.88) gives

$$\Delta E_i^{\text{B,i}} = \frac{Q_i}{r_X^{\text{eff}}} + V_i + \Theta_X. \tag{4.233}$$

Substituting Eqn. (4.24) into this gives

$$\Delta E_i^{\text{B,i}} = \frac{Q_i}{r_X^{\text{eff}}} - a Q_i + a b_X + \Theta_X, \tag{4.234}$$

where we have used Eqn. (4.26). This is equivalent to

$$\Delta E_i^{\text{B,i}} = \left(\frac{1}{r_X^{\text{eff}}} - a \right) Q_i + a b_X + \Theta_X. \tag{4.235}$$

Taking the mean of the above equation over X sites in S gives Eqn. (4.230) after exploiting the linearity of the mean (Eqn. (3.17)). Taking the variance of the above equation over the same set of sites gives Eqn. (4.231) after using Eqn. (4.210). ■

[14] In this chapter, as in Chapter 3, we use the term 'initial state CLS' in place of 'CLS' for the sake of brevity.

4.7.4 Energies

The contributions to the energies E_L, E_M and E from sites in S can also be expressed in terms of $\langle Q\rangle_X^S$ and $\mathrm{Var}(Q)_X^S$ for all X. Denoting these quantities as E_L^S, E_M^S and E^S respectively, we find that

$$E_L^S = \frac{1}{2}a\sum_X |X\cap S|\left[\mathrm{Var}(Q)_X^S + \left(\langle Q\rangle_X^S\right)^2 - 2b_X\langle Q\rangle_X^S + b_X^2\right], \tag{4.236}$$

$$E_M^S = \frac{1}{2}a\sum_X |X\cap S|\left[-\mathrm{Var}(Q)_X^S - \left(\langle Q\rangle_X^S\right)^2 + b_X\langle Q\rangle_X^S\right] \tag{4.237}$$

and

$$E^S = \frac{1}{2}a\sum_X |X\cap S| b_X\left[b_X - \langle Q\rangle_X^S\right]. \tag{4.238}$$

Proof: We will now derive Eqns. (4.236), (4.237) and (4.238), beginning with Eqn. (4.236). Recall that the local energy in the GLCM is given by Eqn. (4.22). The contribution to this from sites in S is

$$E_L^S = \frac{1}{2}a\sum_{i\in S}(Q_i - b_i)^2. \tag{4.239}$$

Splitting the summation in the above equation into contributions from each species, we find that

$$\begin{aligned} E_L^S &= \frac{1}{2}a\sum_X\sum_{i\in X\cap S}(Q_i - b_X)^2 \\ &= \frac{1}{2}a\sum_X\left[\sum_{i\in X\cap S}Q_i^2 + b_X^2\sum_{i\in X\cap S}1 - 2b_X\sum_{i\in X\cap S}Q_i\right] \\ &= \frac{1}{2}a\sum_X\left[|X\cap S|\langle Q_i^2\rangle_{i\in X\cap S} + b_X^2|X\cap S| - 2b_X|X\cap S|\langle Q_i\rangle_{i\in X\cap S}\right], \end{aligned} \tag{4.240}$$

where in the last line we have used the definitions of $\langle Q_i^2\rangle_{i\in X\cap S}$ and $\langle Q_i\rangle_{i\in X\cap S}$:

$$\langle Q_i^2\rangle_{i\in X\cap S} = \frac{1}{|X\cap S|}\sum_{i\in X\cap S}Q_i^2 \tag{4.241}$$

and

$$\langle Q_i\rangle_{i\in X\cap S} = \frac{1}{|X\cap S|}\sum_{i\in X\cap S}Q_i. \tag{4.242}$$

Substituting

$$\langle Q_i^2\rangle_{i\in X\cap S} = \mathrm{Var}(Q)_X^S + \left(\langle Q\rangle_X^S\right)^2 \tag{4.243}$$

into Eqn. (4.240) and simplifying the resulting equation gives Eqn. (4.236), where Eqn. (4.243) follows from the definition of $\mathrm{Var}(Q)_X^S$:

$$\mathrm{Var}(Q)_X^S = \mathrm{Var}(Q)_{i\in X\cap S} = \langle Q_i^2\rangle_{i\in X\cap S} - \Big(\langle Q\rangle_X^S\Big)^2. \tag{4.244}$$

We will now derive Eqn. (4.237). From comparing Eqns. (2.54) and (4.23), it can be seen that the Madelung energy can be expressed as

$$E_M = \frac{1}{2}\sum_i Q_i V_i, \tag{4.245}$$

and hence the contribution to E_M from sites in set S is given by

$$E_M^S = \frac{1}{2}\sum_{i\in S} Q_i V_i. \tag{4.246}$$

This becomes

$$E_M = \frac{1}{2}\sum_{i\in S} Q_i(-aQ_i + ab_i) = \frac{1}{2}\sum_{i\in S}(-aQ_i^2 + ab_iQ_i), \tag{4.247}$$

after using Eqns. (4.24) and (4.26). Splitting the summation in the above into contributions from each species gives

$$\begin{aligned} E_M^S &= \frac{1}{2}\sum_X \sum_{i\in X\cap S}(-aQ_i^2 + ab_X Q_i) \\ &= \frac{1}{2}\sum_X\left[-a\sum_{i\in X\cap S} Q_i^2 + ab_X \sum_{i\in X\cap S} Q_i\right] \\ &= \frac{1}{2}\sum_X\Big[-a|X\cap S|\langle Q_i^2\rangle_{i\in X\cap S} + ab_X|X\cap S|\langle Q\rangle_X^S\Big], \end{aligned} \tag{4.248}$$

where in the last line we have used Eqns. (4.241) and (4.242). Substituting Eqn. (4.243) into this and simplifying the resulting equation gives Eqn. (4.237).

Finally, we will derive Eqn. (4.238). From Eqn. (4.21), it follows that E^S is simply $E_L^S + E_M^S$. Adding Eqns. (4.236) and (4.237) gives

$$E^S = \frac{1}{2}a\sum_X |X\cap S|\Big[-2b_X\langle Q\rangle_X^S + b_X\langle Q\rangle_X^S + b_X^2\Big], \tag{4.249}$$

which becomes Eqn. (4.238) after simplification. ■

4.8 Random Alloys

We will now apply the GLCM to random alloys.

4.8.1 Analytical Expressions

We begin by using the expressions given in the previous section to derive analogous expressions which apply for random alloys.

4.8.1.1 Charges

Using Eqns. (4.202) and (4.203), the following expressions for the mean and variance in Q_i for X sites in random alloys can be derived:

$$\langle Q \rangle_X = aG_0 b_X \tag{4.250}$$

and

$$\mathrm{Var}(Q)_X = (aG_0)^2 \omega j, \tag{4.251}$$

where

$$j = \sum_Y b_Y^2 c_Y \tag{4.252}$$

and

$$\omega = \sum_{\beta=1}^{\infty} g_\beta^2 Z_\beta. \tag{4.253}$$

The values of ω calculated using the g_β obtained from our numerical calculations discussed in Section 4.6 are plotted against aR_{WS} in Fig. 4.4.

Proof: We will now derive Eqns. (4.250) and (4.251). Consider Eqn. (4.250) first. In random alloys,

$$p_\beta^{XY} = c_Y \tag{4.254}$$

for $\beta > 0$, which when substituted into Eqn. (4.202) gives

$$\langle Q \rangle_X = aG_0 \sum_Y b_{YX} c_Y \sum_{\beta=1}^{\infty} g_\beta Z_\beta. \tag{4.255}$$

After using Eqn. (4.84), this becomes

$$\langle Q \rangle_X = -aG_0 \sum_Y b_{YX} c_Y. \tag{4.256}$$

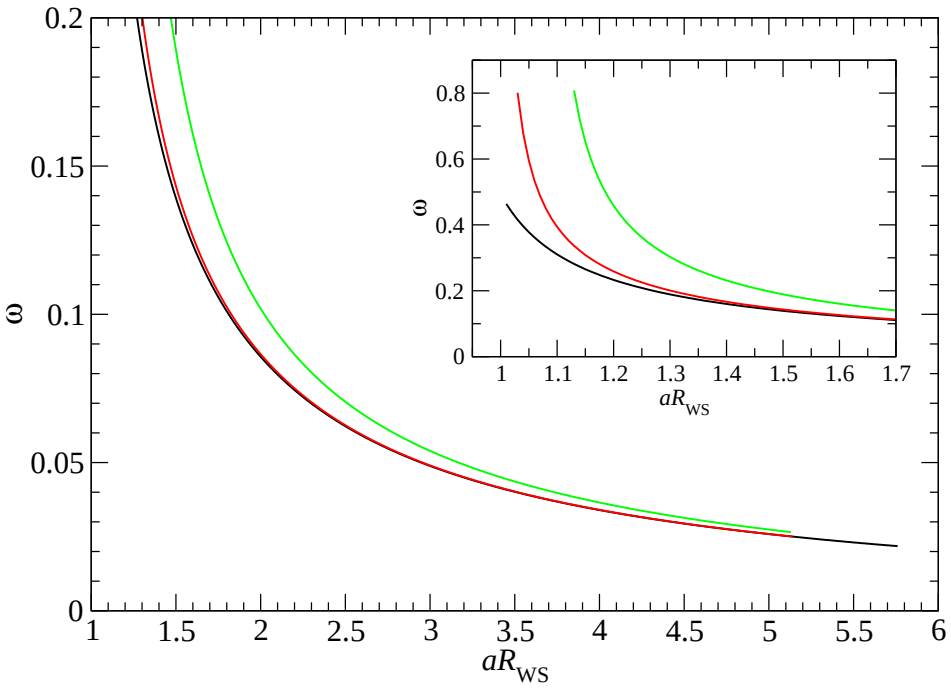


Figure 4.4: ω vs. aR_{WS}. The black, red and green curves correspond to the fcc, bcc and sc lattices respectively. For each value of aR_{WS}, ω has been determined using Eqn. (4.253) and the values of g_β obtained from the numerical calculations described in Section 4.6.

Now,

$$\begin{aligned}\sum_Y b_{YX}c_Y &= \sum_Y (b_Y - b_X)c_Y \\ &= \sum_Y b_Y c_Y - b_X \sum_Y c_Y.\end{aligned} \tag{4.257}$$

where in the first line we have used Eqn. (4.97). The above becomes

$$\sum_Y b_{YX}c_Y = -b_X \tag{4.258}$$

after applying Eqns. (4.47) and (2.51) to the first and second terms respectively. Substituting this into Eqn. (4.256) gives Eqn. (4.250).

Consider now Eqn. (4.251). In random alloys,

$$p_{\beta\gamma}^{XYZ} = c_Y c_Z \tag{4.259}$$

for $\beta, \gamma > 0$. Substituting this and Eqn. (4.254) into Eqn. (4.203) gives

$$\mathrm{Var}(Q)_X = (aG_0)^2 \left\{ \sum_Y b_{YX}^2 c_Y \sum_{\beta=1}^{\infty} g_\beta^2 Z_\beta - \sum_{Y,Z} b_{YX} b_{ZX} c_Y c_Z \sum_{\beta=1}^{\infty} g_\beta^2 Z_\beta. \right\} \tag{4.260}$$

Substituting Eqn. (4.253) into the above and factorising gives

$$\begin{aligned}\mathrm{Var}(Q)_X &= (aG_0)^2 \left[\sum_Y b_{YX}^2 c_Y - \sum_{Y,Z} b_{YX} b_{ZX} c_Y c_Z \right] \omega \\ &= (aG_0)^2 \omega \left[\sum_Y b_{YX}^2 c_Y - \sum_Y b_{YX} c_Y \sum_Z b_{ZX} c_Z \right] \\ &= (aG_0)^2 \omega \left[\sum_Y b_{YX}^2 c_Y - \left(\sum_Y b_{YX} c_Y \right)^2 \right].\end{aligned} \tag{4.261}$$

Now, as can be seen from Eqn. (4.258), the expression in the round brackets equals

$-b_X$. Therefore

$$\begin{aligned}\mathrm{Var}(Q)_X &= (aG_0)^2\omega\left[\sum_Y b_{YX}^2 c_Y - b_X^2\right] \\ &= (aG_0)^2\omega\left[\sum_Y b_{YX}^2 c_Y - \sum_Y b_X^2 c_Y\right] \\ &= (aG_0)^2\omega\left[\sum_Y (b_Y^2 + b_X^2 - 2b_Y b_X)c_Y - \sum_Y b_X^2 c_Y\right],\end{aligned} \tag{4.262}$$

where in the second line we have used Eqn. (2.51), and in the third line we have expanded b_{YX}^2 using Eqn. (4.97). Cancelling the second and final terms in the last line leaves

$$\mathrm{Var}(Q)_X = (aG_0)^2\omega\left[\sum_Y b_Y^2 c_Y - 2b_X\sum_Y b_Y c_Y\right]. \tag{4.263}$$

Noting that the second term in this vanishes on account of Eqn. (4.47) and applying Eqn. (4.252) to the first term gives Eqn. (4.251). ■

There are several points we wish to make regarding Eqns. (4.250) and (4.251). Firstly, as can be seen from comparing Eqns. (4.99) and (4.250), Q_{X0} is the mean charge of an X site in a random alloy - a fact which was stated without proof in Section 4.4. Secondly, in Section 4.2.1 it was pointed out that, if the effective medium approach described in Section 4.1.2 is used in conjunction with the SSCPA effective medium, then b_X is equal to Q_X^{SSCPA} - the charge associated with species X obtained from a SSCPA calculation. Now, Q_X^{SSCPA} purports to be $\langle Q\rangle_X$. However, this is not reflected in Eqn. (4.250): $\langle Q\rangle_X \neq b_X$. This discrepancy stems from the fact that E_M is tacitly and incorrectly assumed to be 0 in SSCPA calculations (see Ref. [24] and references therein for details) - a deficiency which the GLCM does not suffer from. In fact, it was this problem with the SSCPA which prompted the order-N calculations described in Section 2.2.2, and ultimately led to the discovery of the Q-V relations. Our final point regarding Eqns. (4.250) and (4.251) is that the right-hand side of Eqn. (4.251) has no dependence on X, and hence $\mathrm{Var}(Q)_X$ is the same for all species in a random alloy. The same applies to the expression for $\mathrm{Var}(V)_X$ derived in a moment.

4.8.1.2 Madelung Potentials and CLSs

Substituting Eqns. (4.250) and (4.251) into Eqns. (4.227), (4.230), (4.228) and (4.231), and factorising the resulting equations where neccessary, we obtain the following

expressions for the mean and variance in V_i and $\Delta E_i^{\mathrm{B,i}}$ for X sites in random alloys:

$$\langle V\rangle_X = a\Big[1-(aG_0)\Big]b_X, \tag{4.264}$$

$$\langle \Delta E^{\mathrm{B,i}}\rangle_X = \left[a-\left(a-\frac{1}{r_X^{\mathrm{eff}}}\right)(aG_0)\right]b_X+\Theta_X, \tag{4.265}$$

$$\mathrm{Var}(V)_X = a^2(aG_0)^2\omega j \tag{4.266}$$

and

$$\mathrm{Var}(\Delta E^{\mathrm{B,i}})_X = \left(\frac{1}{r_X^{\mathrm{eff}}}-a\right)^2(aG_0)^2\omega j. \tag{4.267}$$

4.8.1.3 Energies

Eqns. (4.250) and (4.251) can also be used to derive expressions for the *intensive* (per atom) local, Madelung and total energies for random alloys.[15] These are

$$E_L = \frac{1}{2}a\Big[(aG_0)^2(\omega+1)-2(aG_0)+1\Big]j \tag{4.268}$$

$$E_M = \frac{1}{2}a\Big[-(aG_0)^2(\omega+1)+(aG_0)\Big]j \tag{4.269}$$

and

$$E = \frac{1}{2}a\Big[1-(aG_0)\Big]j. \tag{4.270}$$

Proof: We will now derive Eqns. (4.268), (4.269) and (4.270). Consider Eqns. (4.236), (4.237) and (4.238). Choosing S to be the set of all sites, in which case the set $X\cap S$ becomes simply X, and dividing through by the total number of sites N in the system under consideration, these equations become

$$E_L = \frac{1}{2}a\sum_X c_X\Big[\mathrm{Var}(Q)_X+\Big(\langle Q\rangle_X\Big)^2-2b_X\langle Q\rangle_X+b_X^2\Big], \tag{4.271}$$

$$E_M = \frac{1}{2}a\sum_X c_X\Big[-\mathrm{Var}(Q)_X-\Big(\langle Q\rangle_X\Big)^2+b_X\langle Q\rangle_X\Big] \tag{4.272}$$

and

$$E = \frac{1}{2}a\sum_X c_X\Big[b_X^2-b_X\langle Q\rangle_X\Big], \tag{4.273}$$

[15] Be aware that we have used the same symbols for the extensive and intensive energies. In this section, only intensive energies are considered.

where we have used the fact that $|X|/N = c_X$. Substituting Eqns. (4.250) and (4.251) into the above gives

$$E_L = \frac{1}{2}a\sum_X c_X\Big[(aG_0)^2\omega j + (aG_0)^2 b_X^2 - 2(aG_0)b_X^2 + b_X^2\Big], \tag{4.274}$$

$$E_M = \frac{1}{2}a\sum_X c_X\Big[-(aG_0)^2\omega j - (aG_0)^2 b_X^2 + (aG_0)b_X^2\Big] \tag{4.275}$$

and

$$E = \frac{1}{2}a\sum_X c_X\Big[b_X^2 - (aG_0)b_X^2\Big]. \tag{4.276}$$

Expanding the summations over X, these become

$$E_L = \frac{1}{2}a\Big[(aG_0)^2\omega j + (aG_0)^2\sum_X b_X^2 c_X - 2(aG_0)\sum_X b_X^2 c_X + \sum_X b_X^2 c_X\Big], \tag{4.277}$$

$$E_M = \frac{1}{2}a\Big[-(aG_0)^2\omega j - (aG_0)^2\sum_X b_X^2 c_X + (aG_0)\sum_X b_X^2 c_X\Big] \tag{4.278}$$

and

$$E = \frac{1}{2}a\Big[\sum_X b_X^2 c_X - (aG_0)\sum_X b_X^2 c_X\Big], \tag{4.279}$$

where we have used Eqn. (2.51). These in turn become Eqns. (4.268), (4.269) and (4.270) respectively after using Eqn. (4.252) and simplifying. ■

4.8.2 Composition Dependence of Physical Quantities

We will now use the above expressions to investigate how the physical properties to which they pertain depend on composition in random ternary alloys. To do this, we will assume that the following quantities are composition-independent: the underlying lattice, a, the electropositivity differences b_{YX} for all Y and X, and the values of r_X^{eff} for all X. Note that, as can be seen from Eqn. (4.232), a consequence of assuming that r_X^{eff} is composition-independent is that Θ_X is also composition-independent. Furthermore, a consequence of assuming that the values of b_{YX} are composition-independent is that the values of b_X are composition-*dependent*. This is required for Eqn. (4.47) to be satisfied for all compositions, and can be seen from Eqn. (4.258): if the values of b_{YX} are fixed, and one changes the values of c_Y, then b_X must change. Table 4.1 provides justification for the assumption that a and the values of b_{YX} are composition-independent. As can be seen from the table, the variation in a_{Cu}, $a_{\text{Zn/Pd}}$ and $(b_{\text{Cu}} - b_{\text{Zn/Pd}})$ with the concentration c of Cu is, at most, $\approx 10\%$, which implies that the accuracy of this

assumption is such that the forthcoming results are at least qualitatively accurate.

Consider a random ternary alloy with a particular composition. Without loss of generality, let us label the three species A, B and C such that $b_A < b_B < b_C$. Our labelling is therefore such that species A is the most electronegative in the alloy, species C is the most electropositive, and species B has an electronegativity/electropositivity between those of species A and C. Using Eqn. (4.97), it can easily be shown that a consequence of this labelling is that the values of b_{CA}, b_{CB} and b_{BA} are all positive, with $b_{CA} > b_{BA}$. Furthermore, using the assumption above that the values of b_{YX} are composition-independent, it can also be shown that, for all compositions, A is the most electronegative species, C is the most electropositive species and B is the species with the intermediate electronegativity/electropositivity.

4.8.2.1 Quantities Proportional to b_A, b_B or b_C

With the above in mind, we will now investigate how b_X varies with composition. As can be seen from Eqns. (4.250), (4.264) and (4.265), this will reveal how $\langle Q \rangle_X$, $\langle V \rangle_X$ and $\langle \Delta E^{\mathrm{B,i}} \rangle_X$ vary with composition. Now, b_A, b_B and b_C can be expressed entirely in terms of the variables c_A, c_B, b_{BA} and b_{CA} as follows:

$$b_A = -b_{BA}c_B - b_{CA}(1 - c_A - c_B), \tag{4.280}$$

$$b_B = b_{BA}c_A - (b_{CA} - b_{BA})(1 - c_A - c_B) \tag{4.281}$$

and

$$b_C = b_{CA}c_A - (b_{BA} - b_{CA})c_B. \tag{4.282}$$

Proof: We will now derive Eqns. (4.280), (4.281) and (4.282). Consider Eqn. (4.258). From this it follows that

$$b_X = -\sum_Y b_{YX}c_Y. \tag{4.283}$$

Applying this to ternary alloys gives

$$b_A = -b_{BA}c_B - b_{CA}c_C, \tag{4.284}$$

$$b_B = -b_{AB}c_A - b_{CB}c_C \tag{4.285}$$

and

$$b_C = -b_{AC}c_A - b_{BC}c_B. \tag{4.286}$$

Substituting

$$c_C = 1 - c_A - c_B, \tag{4.287}$$

which follows from Eqn. (2.51) for ternary alloys, into Eqn. (4.284) gives Eqn. (4.280). Substituting the above equation into Eqn. (4.285) gives

$$b_B = -b_{AB}c_A - b_{CB}(1 - c_A - c_B). \tag{4.288}$$

This becomes Eqn. (4.281) after applying Eqn. (4.110) to b_{AB} and

$$b_{XY} = b_{XZ} - b_{YZ} \tag{4.289}$$

to b_{CB}. The above equation holds for any species X, Y and Z, and its validity can be verified by applying Eqn. (4.97) to each term. Finally, applying Eqn. (4.110) to b_{AC} and the above equation to b_{BC} in Eqn. (4.286) gives Eqn. (4.282). ■

Using Eqns. (4.280), (4.281) and (4.282), we have generated ternary graphs of b_A, b_B and b_C. These are shown in Figs. 4.5, 4.6 and 4.7 respectively. Each of these figures contains three graphs. In all graphs, $b_{CA} = 1$, and either:

- $b_{BA} = 0.2$, which corresponds to species B being more similar, in terms of its electropositivity, to species A than species C;
- $b_{BA} = 0.5$, which corresponds to species B having an electropositivity exactly halfway between those of species A and C;
- $b_{BA} = 0.8$, which corresponds to species B being more similar to species C than species A.

The figures illustrate the following properties of b_A, b_B and b_C, which can be verified analytically using Eqns. (4.280), (4.281) and (4.282):

- b_A takes its maximum value of 0 at $c_A = 1$ (i.e. a 'dilute' random alloy with a vanishing concentration of B and/or C sites). b_A is therefore always non-positive. Furthermore, an increase in the concentration of B or C sites always results in a decrease in b_A, with an increase in c_C having a larger effect than an increase in c_B. The minimum value of b_A is $-b_{CA}$, which occurs at $c_C = 1$.
- b_C takes its minimum value of 0 at $c_C = 1$. b_C is therefore always non-negative. Furthermore, an increase in the concentration of A or B sites always results in an increase in b_C, with an increase in c_A having a larger effect than an increase in c_B. The maximum value of b_C is b_{CA}, which occurs at $c_A = 1$.
- b_B is 0 at $c_B = 1$. Increasing the concentration of A sites always results in b_B being increased; increasing the concentration of C sites always results in b_C being

decreased. Furthermore, the effect of an increase in c_A is smaller than the effect of an increase in c_C if the electropositivity of B is closer to that of A, and larger if the electropositivity of B is closer to that of C. The maximum value of b_B is b_{BA}, which occurs at $c_A = 1$, and the minimum value of b_B is $-b_{CB}$, which occurs at $c_C = 1$.

Since $\langle Q \rangle_X$, $\langle V \rangle_X$ and $\langle \Delta E^{\mathrm{B,i}} \rangle_X$ are proportional to b_X - as can be seen from Eqns. (4.250), (4.264), (4.265) - analogous properties to those just described will therefore apply to these physical quantities. As an example, consider $\langle Q \rangle_A$, the expression for which is given by setting $X = A$ in Eqn. (4.250). Note that, as mentioned earlier, aG_0 is always positive. With this in mind, the properties of b_A described above imply that:

- $\langle Q \rangle_A$ takes its maximum value of 0 at $c_A = 1$, and an increase in the concentration of B or C sites always results in a decrease in $\langle Q \rangle_A$, with an increase in c_C having a larger effect than an increase in c_B.

This makes sense for the following reasons. For $c_A = 1$ almost all sites belong to species A, and hence a vanishing amount of charge transfer occurs. Therefore $\langle Q \rangle_A = 0$ if $c_A = 1$. Increasing the concentration of sites belonging to species B or C, which are both more electropositive than A, results in charge being transferred away from the A sites on average, causing $\langle Q \rangle_A$ to become negative. This effect is more pronounced the higher c_B or c_C is. Furthermore, increasing c_C results in a larger decrease in $\langle Q \rangle_A$ than increasing c_B does since species C is more electropositive than species B, and hence C sites will 'steal' more charge from A sites on average than B sites will.

4.8.2.2 Quantities Proportional to j

Having examined the composition-dependence of b_X in some detail, we will now do the same for the quantity j. As can be seen from Eqns. (4.251), (4.266), (4.267), (4.268), (4.269) and (4.270), this will reveal how $\mathrm{Var}(Q)_X$, $\mathrm{Var}(V)_X$, $\mathrm{Var}(\Delta E^{\mathrm{B,i}})_X$, E_L, E_M and E vary with composition. In a similar manner as was done for b_X above, one can express j entirely in terms of the variables c_A, c_B, b_{BA} and b_{CA}:

$$j = b_{BA}^2 c_A c_B + b_{CA}^2 c_A (1 - c_A - c_B) + (b_{CA} - b_{BA})^2 c_B (1 - c_A - c_B). \tag{4.290}$$

Proof: We will now derive Eqn. (4.290). From Eqns. (4.252) and (2.51), it follows that

$$j = \sum_Y b_Y^2 c_Y \sum_X c_X = \sum_Y \sum_X b_Y^2 c_Y c_X. \tag{4.291}$$

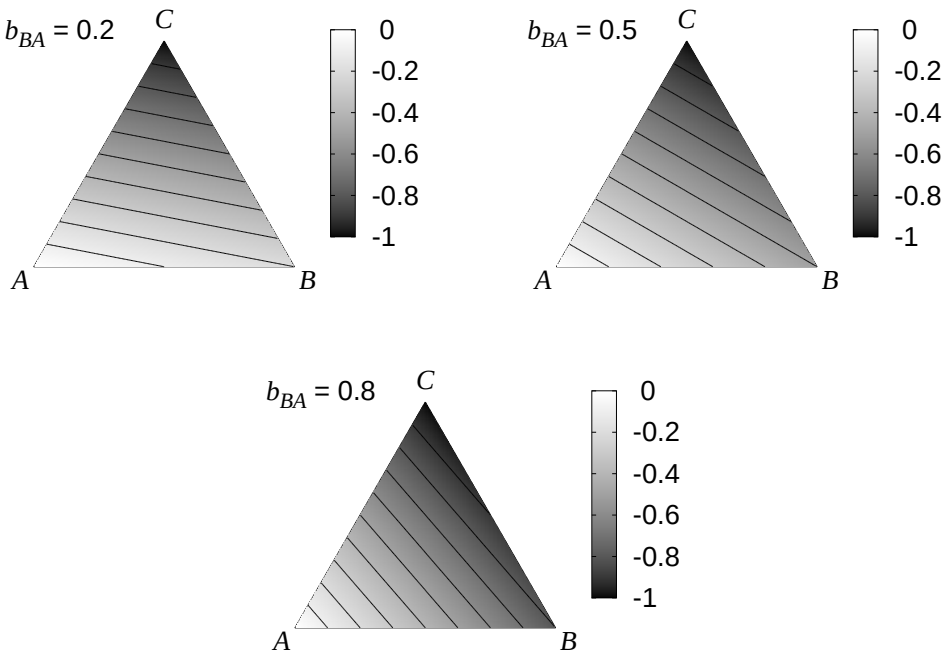


Figure 4.5: Ternary graphs of b_A for $b_{CA} = 1$ and various values of b_{BA}. The corresponding value of b_{BA} for each graph is indicated. In each graph, the contours correspond to curves along which b_A is constant. The contour which is nearest to the bottom left corner of each graph corresponds to $b_A = -0.1$. The change in b_A between adjacent contours is 0.1.

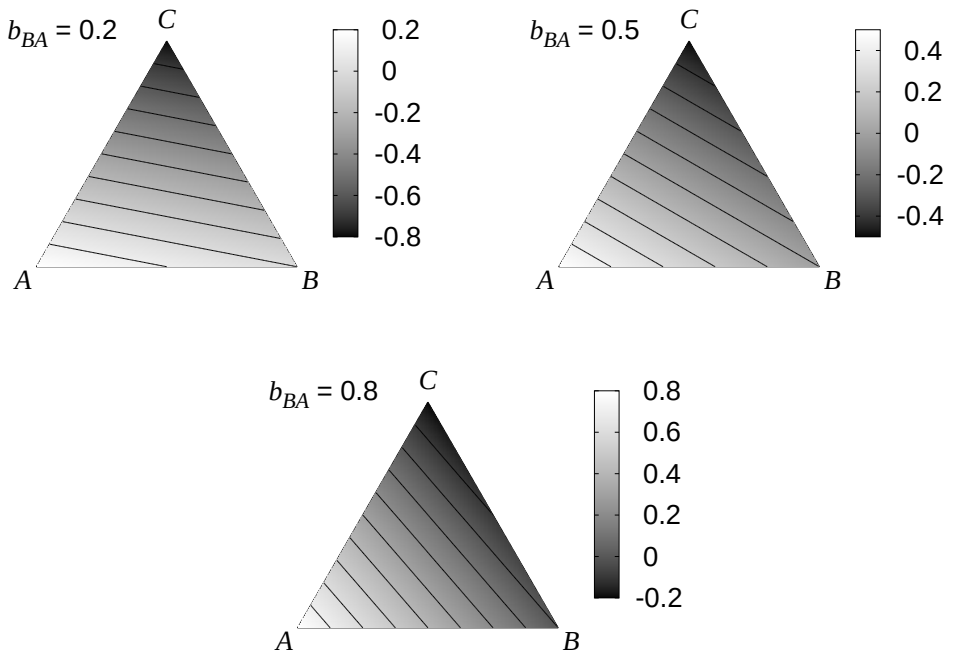


Figure 4.6: Ternary graphs of b_B for $b_{CA} = 1$ and various values of b_{BA}. The corresponding value of b_{BA} for each graph is indicated. In each graph, the contours correspond to curves along which b_B is constant. The contour which ends at the bottom right corner of each graph (i.e. at the point corresponding to $c_B = 1$) corresponds to $b_B = 0$. The change in b_B between adjacent contours is 0.1.

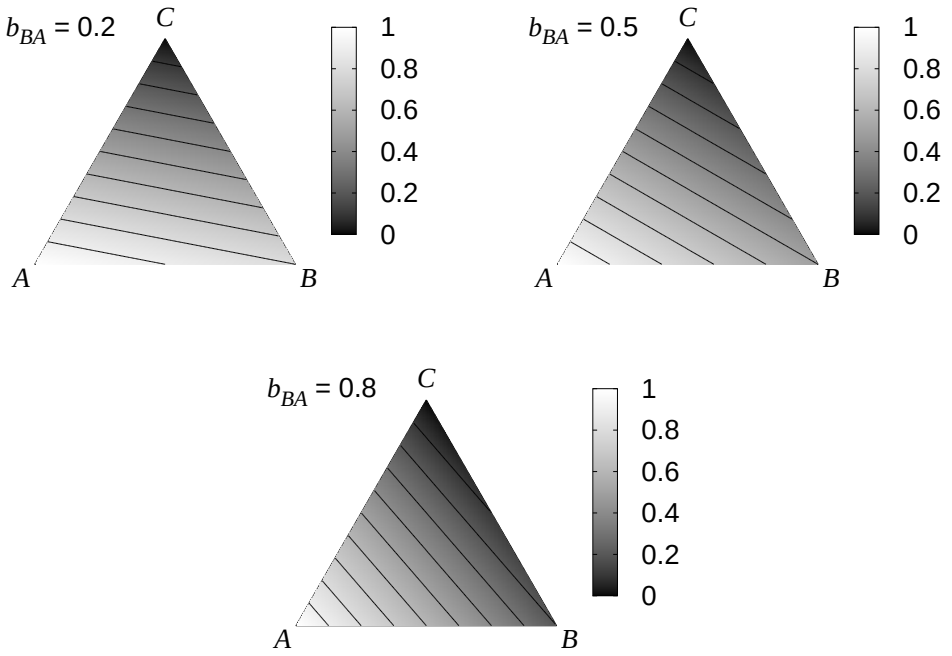


Figure 4.7: Ternary graphs of b_C for $b_{CA} = 1$ and various values of b_{BA}. The corresponding value of b_{BA} for each graph is indicated. In each graph, the contours correspond to curves along which b_C is constant. The contour which is nearest to the top corner of each graph corresponds to $b_C = 0.1$. The change in b_C between adjacent contours is 0.1.

Now,

$$b_{YX}^2 = (b_Y - b_X)^2 = b_Y^2 - 2b_Y b_X + b_X^2, \tag{4.292}$$

where we have used Eqn. (4.97). Rearranging the above for b_Y^2 gives

$$b_Y^2 = b_{YX}^2 + 2b_Y b_X - b_X^2, \tag{4.293}$$

which, when substituted into Eqn. (4.291), gives

$$j = \sum_Y \sum_X (b_{YX}^2 + 2b_Y b_X - b_X^2) c_Y c_X. \tag{4.294}$$

Expanding this, we obtain

$$j = \sum_Y \sum_X b_{YX}^2 c_Y c_X + 2 \sum_Y b_Y c_Y \sum_X b_X c_X - \sum_Y c_Y \sum_X b_X^2 c_X. \tag{4.295}$$

Applying Eqn. (4.47) to the second term and Eqn. (2.51) to the final term, this becomes

$$\begin{aligned} j &= \sum_Y \sum_X b_{YX}^2 c_Y c_X - \sum_X b_X^2 c_X \\ &= \sum_Y \sum_X b_{YX}^2 c_Y c_X - j, \end{aligned} \tag{4.296}$$

where we have used Eqn. (4.252). Solving the above for j gives

$$j = \frac{1}{2} \sum_Y \sum_X b_{YX}^2 c_Y c_X. \tag{4.297}$$

Now, noting that the $Y = A$, $X = B$ term in the double summation has the same value as the $Y = B$, $X = A$ term, and that the $X = Y$ terms vanish since $b_{YY} = 0$, the above can be expressed as a summation over all *combinations* of Y and X as follows:

$$j = \frac{1}{2} \times 2 \sum_{Y,X<Y} b_{YX}^2 c_Y c_X = \sum_{Y,X<Y} b_{YX}^2 c_Y c_X, \tag{4.298}$$

where the notation $Y, X < Y$ indicates that the double summation in the above equation is over all combinations of two species in the alloy, and the factor of 2 in the first equality arises because each combination of Y and $X \neq Y$ occurs in the double summation of Eqn. (4.297) twice. For ternary alloys, the above becomes

$$j = b_{BA}^2 c_B c_A + b_{CA}^2 c_C c_A + b_{CB}^2 c_C c_B. \tag{4.299}$$

Substituting Eqn. (4.287) into this and applying Eqn. (4.289) to b_{CB} gives Eqn. (4.290). ■

Eqn. (4.290) has been used to generate ternary graphs of j, which are shown in Fig. 4.8, and use the same values of b_{BA} and b_{CA} as were considered in Figs. 4.5, 4.6 and 4.7. Fig. 4.8 illustrates the following properties of j, which can be verified analytically using Eqn. (4.290):

- j takes its minimum value of 0 when either $c_A = 1$, $c_B = 1$ or $c_C = 1$. It takes its maximum value of $b_{CA}^2/4$ when $c_A = c_C = 0.5$ (and hence $c_B = 0$).

Since $\mathrm{Var}(Q)_X$, $\mathrm{Var}(V)_X$, $\mathrm{Var}(\Delta E^{\mathrm{B,i}})_X$, E_L, E_M and E are proportional to j - as can be seen from Eqns. (4.251), (4.266), (4.267), (4.268), (4.269) and (4.270) - analogous properties will therefore apply to these physical quantities. As an example, consider $\mathrm{Var}(\Delta E^{\mathrm{B,i}})_X$, the expression for which is given by Eqn. (4.267). Note that, as can be seen from Fig. 4.4, ω is always positive. The same applies to $(1/r_X^{\mathrm{eff}} - a)^2(aG_0)^2$. With this in mind, the properties of j described above imply that:

- $\mathrm{Var}(\Delta E^{\mathrm{B,i}})_X$ takes its minimum value of 0 when either $c_A = 1$, $c_B = 1$ or $c_C = 1$, and takes its maximum value when $c_A = c_C = 0.5$.

Hence the GLCM, in conjunction with the assumptions described earlier, predicts that the initial state disorder broadening in random ternary alloys is maximised at the composition in which the concentrations of the two species with the largest electropositivity difference are equal, and is 0 at the compositions $c_A = 1$, $c_B = 1$ or $c_C = 1$. Note that the fact that there is no disorder broadening if $c_A = 1$, $c_B = 1$ or $c_C = 1$ makes sense since at these compositions there is a vanishing amount of substitutional disorder.

4.9 Summary

To conclude this chapter, we give a summary of our key findings. We began this chapter by deriving the CEFM energy function in order to elucidate the approximations which underpin the CEFM. These approximations were shown to be: the spherical approximation; that the site charges are perturbed from their 'bare' values by only a small amount; and that $E_{L,i}$ - the 'non-Madelung' contribution to the total energy from site i - is a functional only of the contents of site i, and not of the contents of any other site. Three ways in which the last of these approximations can be achieved were highlighted: if outwith site i is assumed to be an effective medium in the evaluation of $E_{L,i}$; if outwith site i is assumed to be a vacuum in the evaluation of $E_{L,i}$; if both the Thomas-Fermi and local density approximations are utilised.

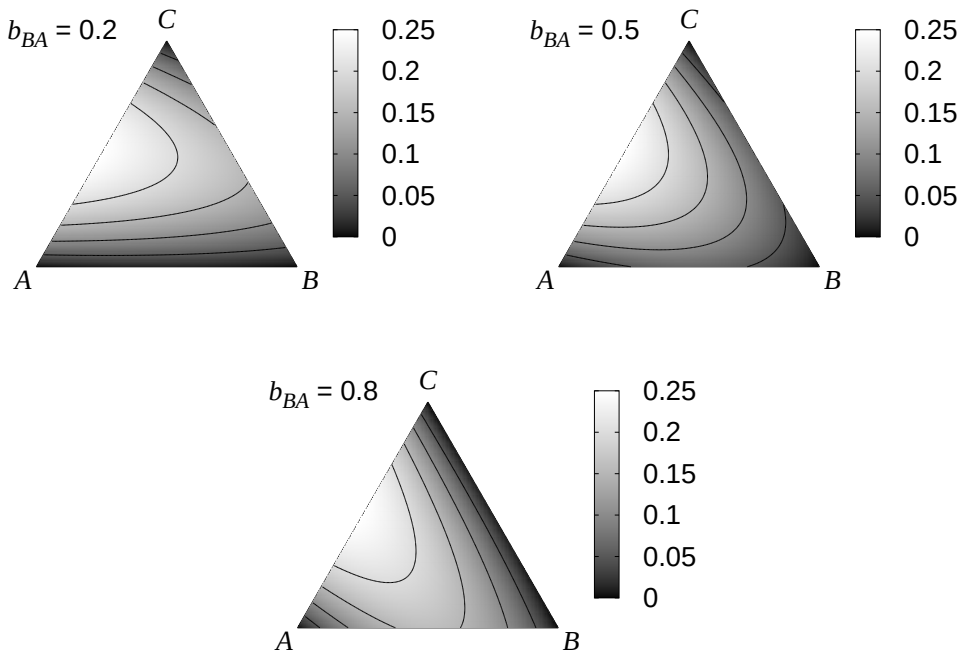


Figure 4.8: Ternary graphs of j for $b_{CA} = 1$ and various values of b_{BA}. The corresponding value of b_{BA} for each graph is indicated. In each graph, the contours correspond to curves along which j is constant. Furthermore, the 'outermost' and 'innermost' visible contours in each graph correspond to $j = 0.05$ and 0.2 respectively. The change in j between adjacent contours is 0.05.

We then considered the GLCM, which is the particular case of the CEFM in which the strength of the 'local interactions' within each site are the same for all sites. The properties of the GLCM were explored in detail. In Section 4.4 expressions for Q_i were derived. These revealed that the values of Q_i in the GLCM can be understood as resulting from charge transfer between all pairs of sites, with the amount of charge transferred between different species depending on their electronegativity differences. It was also shown that the GLCM is equivalent to the OLCM for the case of a binary alloy, and hence the GLCM can be considered to be the generalisation of the OLCM for alloys consisting of more than two species.

In Section 4.6 the 'unknown quantities' in the GLCM were determined for fcc, bcc and sc lattices, and the nature of the screening in the model was explored. An analytical description of the screening was deduced for the limit of weak inter-site Coulomb interactions. Here, the nature of the screening was shown to be universal, i.e. the same for all systems. Numerical calculations were used to determine the nature of the screening away from this limit. It was found that the screening charge distribution around a charge perturbation takes on an increasingly oscillatory form as the strength of the inter-site Coulomb interactions is increased. This was attributed to a Gibb's-phenomenon-like effect arising from the discrete nature of the lattice. Related to this was the observation that the universality in the screening increasingly breaks down at stronger inter-site Coulomb interaction strengths. At the end of Section 4.6 our results were compared to those of Ref. [24], and found to be in quantitative agreement. This was attributed to the fact that all of the approximations which underpin the GLCM are either implicit in the SSLSGF method - which was used in Ref. [24] - or can be justified *a posteriori* from the results of Ref. [24].

In Section 4.7 we used the GLCM to derive analytical expressions for various physical quantities which can be applied to any system. These physical quantities include the mean and variance in Q_i and the initial state CLSs, and the Madelung and total energies of the alloy.

In Section 4.8 analogous expressions were derived for random alloys. These expressions were then used to investigate how the physical quantities to which they pertain vary with composition in a ternary random alloy. In particular, it was observed that the initial state disorder broadening was maximised at the composition where the two species in the alloy with the largest electronegativity difference have equal concentrations, and the remaining species has a vanishing concentration.

Chapter 5

Conclusions and Future Work

In this final chapter, we will describe the key findings of this thesis and discuss possible directions for future work. We begin with our key findings. In Chapter 3 the OLCM was utilised to investigate the relationship between an atom's environment, and its local electronic structure and CLSs, in the bulk and surface regions of disordered alloys with inhomogeneous concentration profiles. The model was found to be a useful tool for rationalising this relationship. Furthermore, systems with inhomogeneous concentration profiles were observed to exhibit interesting phenomena arising from differences in atomic environments, which must be taken into account when interpreting core level XPS spectra. This includes large disorder broadenings, and shifts in the work function induced by differences in the surface and bulk concentrations. The motivation behind Chapter 4 was to develop a model which did not suffer from the shortcomings of the OLCM, but retained its conceptual simplicity, and hence - like the OLCM - can be used as a simple framework for the interpretation of experimental results and the rationalisation of complicated phenomena. Such a model was found in the GLCM, which is a particular case of the CEFM, and - unlike the OLCM - can treat alloys containing any number of species. With the GLCM, the nature of screening in disordered alloys was examined in detail, as was the concentration-dependence of various physical quantities in ternary random alloys.

The most obvious manner in which to proceed is to apply the GLCM to specific systems. This would both test the accuracy of the model, and add insight into the nature of the local electronic structure within the systems under consideration. It may be interesting to investigate systems of the type considered in Chapter 3, i.e. systems with surfaces and/or layer-dependent concentrations, though consisting of more than two species. Such systems were investigated in the recent joint theoretical and experimental studies of Refs. [69] and [70].

The strength of the GLCM is its simplicity: it allows analytical expressions to be

derived for various physical quantities. However, the physical quantities which can be investigated using the GLCM are by no means limited to those considered in Chapter 4. In this chapter, expressions pertaining to the initial state CLSs are given. One could also derive analogous expressions for the *total* CLSs. Recall that the total CLS ΔE_i^{B} of core level i is the sum of the initial state contribution $\Delta E_i^{\mathrm{B,i}}$ and the final state contribution $\Delta E_i^{\mathrm{B,f}}$. Recall also that, in the final state pertaining to core level i, there is a core hole within site i which results in a redistribution (relative to the initial state) of electrons throughout the system. One could represent the final state within the GLCM by changing the value of b_i from b_X to b_X^*, where b_X^* is the bare charge of an X site with an ionised atomic core, and X is the species of site i. The $z+1$ approximation described in Section 2.3.3.1 could be used in order to determine b_X^*. Within this approximation, b_X^* is the bare charge of a site belonging to the element above X in the periodic table. With the above in mind, one could derive an expression for ΔE_i^{B} by substituting the initial and final state energies, calculated using Eqns. (4.21), (4.22) and (4.23), into Eqn. (2.75).

The aforementioned method for determining ΔE_i^{B} utilises the total energies method described in Section 2.3.3.1. An alternative approach is to use the potential model, which has the benefit of allowing the initial and final state contributions to be separated - which is not possible in the total energies method. In fact, we have used the potential model throughout this thesis in order to calculate $\Delta E_i^{\mathrm{B,i}}$. To determine $\Delta E_i^{\mathrm{B,f}}$ via the potential model, the relevant equation is Eqn. (2.85), from which it is clear that one must know how the electrostatic potential V_i^{tot} at nucleus i varies with the occupancy η of the core level within site i. In Chapter 4 the nature of the screening of a charge perturbation - such as that arising due to a change in the occupancy of a core level - was explored in detail. The results of this investigation could be used to deduce how V_i^{tot} varies with η, and hence $\Delta E_i^{\mathrm{B,f}}$. However, in order to do this, one must know how the valence charge is distributed *within* site i at each value of η. Unfortunately, this information is not provided by the GLCM, which only specifies the total amount of charge on each site, and not how this charge is distributed within each site. The same applies for the OLCM. Because of this, it has been necessary in this thesis to make some assumptions regarding how the valence charge is distributed within each site. In Chapter 3 we assumed that the valence charge within site i was located at an effective radius of $r_i^{\mathrm{eff}} = R_1/2$ for all sites; in Chapter 4 we assumed that r_i^{eff} takes the same value r_X^{eff} for all X sites. Both of these assumptions were applied only to the initial state, and ignore the possibility that r_i^{eff} depends on the environment of site i. This is not the case in reality, though the extent to which it is a valid approximation in the initial state requires further investigation. A more justifiable approach is to decompose the valence charge density within each site into components associated with each angular

momentum quantum number l,[1] and to assume that the effective radius $r_i^{\text{eff},l}$ associated with each l-component is the same for all sites. Note that this does not eliminate the aforementioned problem: $r_i^{\text{eff},l}$ will still in general depend on environment; however, it is expected that, for sites belonging to the same species, the variation in the values of $r_i^{\text{eff},l}$ for a given l will be less than the variation in the values of r_i^{eff}. The analogous equation to Eqn. (2.86) in this case is

$$V_i^{\text{tot}} = \sum_l \frac{Q_i^{\text{val},l}}{r_i^{\text{eff},l}} + V_i^{\text{core}} + V_i, \tag{5.1}$$

where $Q_i^{\text{val},l}$ is the valence charge within site i associated with angular momentum l. Such a decomposition has been shown to be essential in some systems when calculating $\Delta E_i^{\text{B,f}}$ - as is summarised in Ref. [27]. This should be borne in mind when calculating $\Delta E_i^{\text{B,f}}$ via the potential model as described above.

Another physical quantity which could be examined with the GLCM is the alloy-metal *Auger kinetic energy shift*. Some time after the creation of a core hole of type t through photoemission, an electron initially occupying a core level of type u may de-excite to fill the hole in the t core level while an electron in a core level of type v is simultaneously ejected from the system. This is known as a *t-u-v* core-core-core *Auger transition*. If all core levels are bound to site i, the kinetic energy of the ejected *Auger electron* can be shown to be

$$E_{tuv,i}^{\text{K,Auger}} = E_{t,i}^{\text{B}} - E_{uv,i}^{\text{B}}, \tag{5.2}$$

where $E_{t,i}^{\text{B}}$ is the binding energy of the t core level, and $E_{uv,i}^{\text{B}}$ is the two-electron binding energy associated with the simultaneous removal of two electrons: one from a u core level and one from a v core level. One can define the Auger kinetic energy shift for the Auger transition associated with site i analogously to ΔE_i^{B} in Eqn. (2.74):

$$\Delta E_{tuv,i}^{\text{K,Auger}} = E_{tuv,i}^{\text{K,Auger}} - E_{tuv,\text{ref}}^{\text{K,Auger}}, \tag{5.3}$$

where $E_{tuv,\text{ref}}^{\text{K,Auger}}$ is the Auger kinetic energy for the *t-u-v* transition associated with any site in a pure X metal, and X is the species of site i. Using Eqn. (5.2), it can be shown that

$$\Delta E_{tuv,i}^{\text{K,Auger}} = \Delta E_{t,i}^{\text{B}} - \Delta E_{uv,i}^{\text{B}}. \tag{5.4}$$

Earlier in this chapter we outlined how to derive an expression for $\Delta E_{t,i}^{\text{B}}$ using the

[1] Within the spherical approximation, the spin orbitals within each site are eigenstates of the angular momentum operator with respect to the site's nucleus, and can be classified according to the angular momentum quantum number l. See, for instance, Ref. [1] for details.

total energies method. An analogous expression would also apply to $\Delta E^{\mathrm{B}}_{uv,i}$, except that in this case b^{*}_{X} is replaced by b^{**}_{X} - the bare charge of an X site with a doubly, not singly, ionised atomic core. With such expressions for $\Delta E^{\mathrm{B}}_{t,i}$ and $\Delta E^{\mathrm{B}}_{uv,i}$, another for $\Delta E^{\mathrm{K,Auger}}_{tuv,i}$ could be derived by using the above equation. With this, one could investigate the disorder broadening in the Auger electron spectrum, which, similarly to the photoelectron spectrum, provides information regarding the distribution of environments within the system under consideration. While disorder broadening of Auger spectra has not yet conclusively been observed experimentally, [71, 72] preliminary model [73] and *ab initio* [38] investigations have already been performed.

To conclude, we wish to highlight the fact that there are some effects which cannot be accounted for by the GLCM. Throughout this thesis we have used the spherical approximation - which was described at the beginning of Section 2.2. One must go beyond this approximation in order to obtain a quantitative description of the electron density within disordered alloys [25, 74].[2] This is especially true for regions near surfaces. A generalisation of the CEFM has been described in Ref. [33] which does not rely upon the spherical approximation. While the GLCM could be generalised in an analogous manner, it is not clear whether this would be fruitful. The strength of the GLCM over the CEFM is its simplicity, with which comes a small loss in accuracy relative to the CEFM. It is not clear whether the gain in accuracy achieved by generalising the GLCM to go beyond the spherical approximation is worth the resulting loss in simplicity. As well as the validity of the spherical approximation, we have also assumed throughout this thesis that the nuclei of the system under consideration form an undistorted crystal lattice. The breakdown of this assumption can have far-reaching consequences. As was mentioned in Section 2.3.4, the addition of distortions to the crystal lattice of CuAu results in a reversal of the average relationship between a site's CLS and its number of unlike nearest neighbours. The reasons for this are not known, and warrant further investigation. The GLCM, suitably modified to treat lattice distortions, may add insight into this phenomenon; though it would be optimistic to expect that anything more than a qualitative understanding could be achieved.

[2] Any claims of quantitative accuracy which we have made in this thesis have been made with reference to *ab initio* calculations utilising approximations analogous to the spherical approximation.

Bibliography

[1] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, first ed., 2004.

[2] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, first ed., 1999.

[3] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 1996.

[4] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Saunders College, first ed., 1976.

[5] G. Grosso and G. P. Parravicini, *Solid State Physics*. Academic Press, 2000.

[6] H. F. Schaefer, *Quantum Chemistry: The Development of Ab Initio Methods in Molecular Electronic Structure Theory*. Clarendon Press, 1984.

[7] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.

[8] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, pp. A1133–A1138, Nov 1965.

[9] A. Ruzsinszky and J. P. Perdew, "Twelve outstanding problems in ground-state density functional theory: A bouquet of puzzles," *Computational and Theoretical Chemistry*, vol. 963, no. 1, pp. 2–6, 2011.

[10] J. S. Faulkner, "The modern theory of alloys," *Progress in Materials Science*, vol. 27, no. 1-2, pp. 1–187, 1982.

[11] H. Ebert, D. Kdderitzsch, and J. Minr, "Calculating condensed matter properties using the KKR-Green's function method - recent developments and applications," *Reports on Progress in Physics*, vol. 74, no. 9, p. 096501, 2011.

[12] J. Zabloudil, *Electron Scattering in Solid Matter: A Theoretical and Computational Treatise.* Springer series in solid-state sciences: 147, Springer, 2005.

[13] W. Kohn, "Nobel lecture: Electronic structure of matter-wave functions and density functionals," *Rev. Mod. Phys.*, vol. 71, pp. 1253–1266, Oct 1999.

[14] J. F. Janak, "Proof that $\partial \mathrm{E}/\partial n_i = \epsilon$ in density-functional theory," *Phys. Rev. B*, vol. 18, pp. 7165–7168, Dec 1978.

[15] P. Soven, "Coherent-potential model of substitutional disordered alloys," *Phys. Rev.*, vol. 156, pp. 809–813, Apr 1967.

[16] D. D. Johnson, D. M. Nicholson, F. J. Pinski, B. L. Gyorffy, and G. M. Stocks, "Density-functional theory for random alloys: Total energy within the coherent-potential approximation," *Phys. Rev. Lett.*, vol. 56, pp. 2088–2091, May 1986.

[17] D. D. Johnson, D. M. Nicholson, F. J. Pinski, B. L. Györffy, and G. M. Stocks, "Total-energy and pressure calculations for random substitutional alloys," *Phys. Rev. B*, vol. 41, pp. 9701–9716, May 1990.

[18] E. N. Economou, *Green's Functions in Quantum Physics.* Springer series in solid-state sciences: 7, Springer, 2006.

[19] Y. Wang, G. M. Stocks, W. A. Shelton, D. M. C. Nicholson, Z. Szotek, and W. M. Temmerman, "Order-N multiple scattering approach to electronic structure calculations," *Phys. Rev. Lett.*, vol. 75, pp. 2867–2870, Oct 1995.

[20] I. A. Abrikosov, A. M. N. Niklasson, S. I. Simak, B. Johansson, A. V. Ruban, and H. L. Skriver, "Order-N Green's function technique for local environment effects in alloys," *Phys. Rev. Lett.*, vol. 76, pp. 4203–4206, May 1996.

[21] I. A. Abrikosov, S. I. Simak, B. Johansson, A. V. Ruban, and H. L. Skriver, "Locally self-consistent Green's function approach to the electronic structure problem," *Phys. Rev. B*, vol. 56, pp. 9319–9334, Oct 1997.

[22] J. S. Faulkner, Y. Wang, and G. M. Stocks, "Electrons in extended systems," *Phys. Rev. B*, vol. 52, pp. 17106–17111, Dec 1995.

[23] J. S. Faulkner, Y. Wang, and G. M. Stocks, "Coulomb energies in alloys," *Phys. Rev. B*, vol. 55, pp. 7492–7507, Mar 1997.

[24] A. V. Ruban and H. L. Skriver, "Screened Coulomb interactions in metallic alloys. I. Universal screening in the atomic-sphere approximation," *Phys. Rev. B*, vol. 66, p. 024201, Jun 2002.

[25] F. J. Pinski, "Charge fluctuations in alloys: A coarse-grained model," *Phys. Rev. B*, vol. 57, pp. 15140–15143, Jun 1998.

[26] T. L. Underwood, P. D. Lane, N. Miller, R. Stoker, and R. J. Cole, "Charge models for electron spectroscopy of disordered alloys," *Phys. Rev. B*, vol. 79, p. 024203, Jan 2009.

[27] R. Cole and P. Weightman, "Disorder broadening of core levels: Insights into alloy electronic structure," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 178-179, pp. 112–122, 2010.

[28] R. Magri, S.-H. Wei, and A. Zunger, "Ground-state structures and the random-state energy of the madelung lattice," *Phys. Rev. B*, vol. 42, pp. 11388–11391, Dec 1990.

[29] C. Wolverton, A. Zunger, S. Froyen, and S.-H. Wei, "Point-charge electrostatics in disordered alloys," *Phys. Rev. B*, vol. 54, pp. 7843–7856, Sep 1996.

[30] E. Bruno and L. Zingales, "Are random metallic alloys charge glasses?," *Philosophical Magazine*, vol. 84, no. 13-16, pp. 1621–1629, 2004.

[31] E. Bruno, L. Zingales, and Y. Wang, "Charge distributions in metallic alloys: A charge-excess functional theory approach," *Phys. Rev. Lett.*, vol. 91, p. 166401, Oct 2003.

[32] V. Drchal, R. Hammerling, and P. Weinberger, "Screening, charge distribution, and electron correlations in metallic alloys," *Phys. Rev. B*, vol. 74, p. 214202, Dec 2006.

[33] E. Bruno, F. Mammano, A. Fiorino, and E. V. Morabito, "Coarse-grained density functional theories for metallic alloys: Generalized coherent-potential approximations and charge-excess functional theory," *Phys. Rev. B*, vol. 77, p. 155108, Apr 2008.

[34] C. R. Brundle and A. D. Baker, *Electron Spectroscopy: Theory, Techniques and Applications.* Academic Press, 1977.

[35] U. Gelius, "Binding energies and chemical shifts in ESCA," *Physica Scripta*, vol. 9, no. 3, p. 133, 1974.

[36] S. Doniach and M. Sunjic, "Many-electron singularity in X-ray photoemission and X-ray line spectra from metals," *Journal of Physics C: Solid State Physics*, vol. 3, no. 2, p. 285, 1970.

[37] B. Johansson and N. Mårtensson, "Core-level binding-energy shifts for the metallic elements," *Phys. Rev. B*, vol. 21, pp. 4427–4457, May 1980.

[38] W. Olovsson, T. Marten, E. Holmstrm, B. Johansson, and I. A. Abrikosov, "First principle calculations of core-level binding energy and Auger kinetic energy shifts in metallic solids," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 178-179, pp. 88–99, 2010.

[39] D. Shirley and C. Fadley, "X-ray photoelectron spectroscopy in north america the early years," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 137-140, pp. 43–58, 2004.

[40] A. R. Williams and N. D. Lang, "Core-level binding-energy shifts in metals," *Phys. Rev. Lett.*, vol. 40, pp. 954–957, Apr 1978.

[41] M. Weinert and R. E. Watson, "Core-level shifts in bulk alloys and surface adlayers," *Phys. Rev. B*, vol. 51, pp. 17168–17180, Jun 1995.

[42] R. J. Cole, N. J. Brooks, and P. Weightman, "Madelung potentials and disorder broadening of core photoemission spectra in random alloys," *Phys. Rev. Lett.*, vol. 78, pp. 3777–3780, May 1997.

[43] R. J. Cole and P. Weightman, "Electrostatics in disordered Cu_xPd_{1-x} alloys," *Journal of Physics: Condensed Matter*, vol. 10, no. 25, p. 5679, 1998.

[44] D. Lewis, R. J. Cole, and P. Weightman, "Observation of disorder broadening of core photoelectron spectra of CuZn alloys," *Journal of Physics: Condensed Matter*, vol. 11, no. 43, p. 8431, 1999.

[45] A. Newton, A. Vaughan, R. Cole, and P. Weightman, "Disorder broadening of core level photoemission spectra in Cu_xPt_{1-x} alloys," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 107, no. 2, pp. 185–191, 2000.

[46] A. Newton, S. Haines, P. Weightman, and R. Cole, "Disorder induced core photoelectron linewidth broadening in AgPd alloys," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 136, no. 3, pp. 235–238, 2004.

[47] T. Marten, I. A. Abrikosov, W. Olovsson, B. Johansson, R. J. Cole, G. Beamson, S. R. Haines, and P. Weightman, "Suppression of disorder broadening of core-level photoelectron lines in CuAu alloys by inhomogeneous lattice distortion," *Phys. Rev. B*, vol. 79, p. 012201, Jan 2009.

[48] J. S. Faulkner, Y. Wang, and G. M. Stocks, "Core level chemical shifts in metallic alloys," *Phys. Rev. Lett.*, vol. 81, pp. 1905–1908, Aug 1998.

[49] T. Marten, W. Olovsson, S. I. Simak, and I. A. Abrikosov, "*Ab initio* study of disorder broadening of core photoemission spectra in random $Cu - Pd$ and $Ag - Pd$ alloys," *Phys. Rev. B*, vol. 72, p. 054210, Aug 2005.

[50] W. Olovsson, E. Holmström, T. Marten, I. A. Abrikosov, and A. M. N. Niklasson, "Interface core-level shifts as a probe of embedded thin-film quality," *Phys. Rev. B*, vol. 84, p. 085431, Aug 2011.

[51] P. P. Ewald, "Die berechnung optischer und elektrostatischer gitterpotentiale," *Annalen der Physik*, vol. 369, no. 3, pp. 253–287, 1921.

[52] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, "Iterative minimization techniques for *ab initio* total-energy calculations: molecular dynamics and conjugate gradients," *Rev. Mod. Phys.*, vol. 64, pp. 1045–1097, Oct 1992.

[53] O. V. Bystrenko, "Calculation of the Coulomb energy in quasi-two-dimensional systems," *Phys. Rev. E*, vol. 65, p. 037702, Mar 2002.

[54] R. J. Cole and P. Weightman, "Madelung potentials in disordered systems," *Journal of Physics: Condensed Matter*, vol. 9, no. 26, p. 5609, 1997.

[55] P. Weightman, "X-ray-excited Auger and photoelectron spectroscopy," *Reports on Progress in Physics*, vol. 45, no. 7, p. 753, 1982.

[56] D. Zwillinger and S. Kokoska, *CRC Standard Probability and Statistics Tables and Formulae*. Chapman and Hall/CRC, 2000.

[57] V. Medicherla and W. Drube, "Electronic structure of PdAg (100) ordered surface alloys using synchrotron radiation," *Applied Surface Science*, vol. 256, no. 2, pp. 376–379, 2009.

[58] E. Bruno, L. Zingales, and A. Milici, "Local charge excesses in metallic alloys: A local-field coherent potential approximation theory," *Phys. Rev. B*, vol. 66, p. 245107, Dec 2002.

[59] A. Zangwill, *Physics at Surfaces*. Cambridge University Press, first ed., 1988.

[60] U. Bardi, "The atomic structure of alloy surfaces and surface alloys," *Reports on Progress in Physics*, vol. 57, no. 10, p. 939, 1994.

[61] R. Monnier, "First-principles approaches to surface segregation," *Philosophical Magazine Part B*, vol. 75, no. 1, pp. 67–144, 1997.

[62] A. Ruban and H. Skriver, "Calculated surface segregation in transition metal alloys," *Computational Materials Science*, vol. 15, no. 2, pp. 119–143, 1999.

[63] E. Holmström, W. Olovsson, I. A. Abrikosov, A. M. N. Niklasson, B. Johansson, M. Gorgoi, O. Karis, S. Svensson, F. Schäfers, W. Braun, G. Öhrwall, G. Andersson, M. Marcellini, and W. Eberhardt, "Sample preserving deep interface characterization technique," *Phys. Rev. Lett.*, vol. 97, p. 266106, Dec 2006.

[64] E. Bruno, "The charge excess functional theory and ordering properties of metallic alloys," *Materials Science and Engineering: A*, vol. 462, no. 1-2, pp. 456–459, 2007.

[65] H. Ltkepohl, *Handbook of Matrices*. Wiley, 1996.

[66] A. D. Polyanin, *Handbook of Linear Partial Differential Equations for Engineers and Scientists*. Chapman and Hall/CRC, 2002.

[67] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*. Society for Industrial and Applied Mathematics, third ed., 1999.

[68] J. M. Cowley, "Short- and long-range order parameters in disordered solid solutions," *Phys. Rev.*, vol. 120, pp. 1648–1657, Dec 1960.

[69] W. Olovsson, L. Bech, T. H. Andersen, Z. Li, S. V. Hoffmann, B. Johansson, I. A. Abrikosov, and J. Onsgaard, "Core-level shifts for two- and three-dimensional bimetallic Pd_xCu_{1-x} and Pd_xAg_{1-x} alloys on Ru(0001)," *Phys. Rev. B*, vol. 72, p. 075444, Aug 2005.

[70] S. Granroth, R. Knut, M. Marcellini, G. Andersson, S. Svensson, O. Karis, M. Gorgoi, F. Schäfers, W. Braun, W. Eberhardt, W. Olovsson, E. Holmström, and N. Mårtensson, "Investigation of interface properties of Ni/Cu multilayers by high kinetic energy photoelectron spectroscopy," *Phys. Rev. B*, vol. 80, p. 094104, Sep 2009.

[71] Z.-T. Jiang, S. M. Thurgate, G. van Riessen, P. Wilkie, and C. Creagh, "The Ag $M_5N_{45}N_{45}$ Auger photoelectron coincidence spectra of disordered $Ag_{0.5}Pd_{0.5}$ alloy," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 130, no. 1-3, pp. 33–41, 2003.

[72] M. Ohno, "The origin of the Auger-photoelectron coincidence spectroscopy (APECS) spectral line broadening of AgPd alloy," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 149, no. 1-3, pp. 1–5, 2005.

[73] R. Stoker, M. Szmigiel, N. Miller, and R. Cole, "Disorder broadening of alloy Auger spectra," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 162, no. 3, pp. 127–133, 2008.

[74] A. V. Ruban, S. I. Simak, P. A. Korzhavyi, and H. L. Skriver, "Screened Coulomb interactions in metallic alloys. II. Screening beyond the single-site and atomic-sphere approximations," *Phys. Rev. B*, vol. 66, p. 024202, Jun 2002.

Publications

T.L. Underwood and R.J. Cole, "Core level broadening at the surfaces of disordered systems," *Journal of Electron Spectroscopy and Related Phenomena*, vol. 181, no. 2-3, pp. 220-224, 2009.